

Db 84 AGCGGACTTCCAGTGTCTTCAGCAGGCGCGCGCGGGGCTGGCGGCACCTTTCCGGGC 143
Qy 282 CTTCTGTGTGTGGGCTGAGGACTTCTACAGATCTGTTCGGCGCGCGCGGACCGG 341
Db 144 CTTCTGTGTGTGGGCTGAGGACTTCTACAGATCTGTTCGGCGCGCGGACCGG 203
Qy 342 GGTGCGGCTGTCAACTCCTAGGAGCAGGCTCTTCCAGCTGGGAGGCTTATTCTC 401
Db 204 GGTGCGGCTGTCAACTCCTAGGAGCAGGCTCTTCCAGCTGGGAGGCTTATTCTC 263
Qy 402 GGGCTCGAGGGGCGAGCTGAAGCCCGGGGCGCGCATCTTCTTTTCAGCGCAGAGATGT 461
Db 264 GGGCTCGAGGGGCGAGCTGAAGCCCGGGGCGCGCATCTTCTTTTCAGCGCAGAGATGT 323
Qy 462 CTTGAGCACCCTGCTGGCCCGGGAAGAGCGTGTGGCA CGGCTTCGACCCGCGGGCG 521
Db 324 CTTGAGCACCCTGCTGGCCCGGGAAGAGCGTGTGGCA CGGCTTCGACCCGCGGGCG 383
Qy 522 CCGCTGACGACGAGCTACTGCGAGCGTGGCGAGCGGAGGCGCGCGCGCGCGGCA 581
Db 384 CCGCTGACGACGAGCTACTGCGAGCGTGGCGAGCGGAGGCGCGCGCGCGCGGCA 443
Qy 582 GGGCTGTCTGCTGTGGCGGCGAGGCTGTGGAGCAGGAGCGCGGAGCTTCTTCCAAAG 641
Db 444 GGGCTGTCTGCTGTGGCGGCGAGGCTGTGGAGCAGGAGCGCGGAGCTTCTTCCAAAG 503
Qy 642 CTTGCTGTCTGTGATCGAAGCAGCGTCATGACCTCTTCTTCCAAAG 690
Db 504 CTTGCTGTCTGTGATCGAAGCAGCGTCATGACCTCTTCTTCCAAAG 552

RESULT 2

US-10-131-241-50
; Sequence 50, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; FILE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10131.241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Canine sp.
US-10-131-241-50

Query Match 63.8%; Score 529; DB 14; Length 552;
Best Local Similarity 100.0%; Pred. No. 5.3e-255;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 162 GGTCTGCTGCTGTGGGCTTGAACAGCCCGCGGCGGCGGATGCGAGGATCCGGG 221
Db 24 GGTCTGCTGCTGTGGGCTTGAACAGCCCGCGGCGGCGGATGCGAGGATCCGGG 83
Qy 222 AGCGGACTTCCAGTGTCTTCAGAGGCGCGCGCGGCGGCTGGCGGCGGCTTCCGGGC 281
Db 84 AGCGGACTTCCAGTGTCTTCAGAGGCGCGCGCGGCGGCTGGCGGCGGCTTCCGGGC 143
Qy 282 CTTCTGTGTGTGGGCTGAGGACTTCTACAGATCTGTTCGGCGCGCGGACCGG 341
Db 144 CTTCTGTGTGTGGGCTGAGGACTTCTACAGATCTGTTCGGCGCGCGGACCGG 203

Qy 342 GGTGCGGCTGTCAACTCCTAGGAGCAGGCTCTTCCCGCAGCTGGGAGGCTTATTCTC 401
Db 204 GGTGCGGCTGTCAACTCCTAGGAGCAGGCTCTTCCCGCAGCTGGGAGGCTTATTCTC 263
Qy 402 GGGCTCGAGGGGCGAGCTGAAGCCCGGGGCGCGCATCTTCTTTTCAGCGCAGAGATGT 461
Db 264 GGGCTCGAGGGGCGAGCTGAAGCCCGGGGCGCGCATCTTCTTTTCAGCGCAGAGATGT 323
Qy 462 CTTGAGCACCCTGCTGGCCCGGGAAGAGCGTGTGGCA CGGCTTCGACCCGCGGGCG 521
Db 324 CTTGAGCACCCTGCTGGCCCGGGAAGAGCGTGTGGCA CGGCTTCGACCCGCGGGCG 383
Qy 522 CCGCTGACGACGAGCTACTGCGAGCGTGGCGAGCGGAGGCGCGCGCGCGGCA 581
Db 384 CCGCTGACGACGAGCTACTGCGAGCGTGGCGAGCGGAGGCGCGCGCGCGGCA 443
Qy 582 GGGCTGTCTGCTGTGGCGGCGAGGCTGTGGAGCAGGAGCGCGGAGCTTCTTCCAAAG 641
Db 444 GGGCTGTCTGCTGTGGCGGCGAGGCTGTGGAGCAGGAGCGCGGAGCTTCTTCCAAAG 503
Qy 642 CTTGCTGTCTGTGATCGAAGCAGCGTCATGACCTCTTCTTCCAAAG 690
Db 504 CTTGCTGTCTGTGATCGAAGCAGCGTCATGACCTCTTCTTCCAAAG 552

RESULT 3

US-10-131-241-51
; Sequence 51, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; FILE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10131.241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
; LENGTH: 632
; TYPE: DNA
; ORGANISM: Murinae sp.
US-10-131-241-51

Query Match 47.3%; Score 392; DB 14; Length 632;
Best Local Similarity 100.0%; Pred. No. 2.2e-186;
Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 299 TGCAGGACCTCTACAGCATCTGCGCGCGCGCGCGCGGAGGCGCGGCTCGTCAACC 358
Db 241 TGCAGGACCTCTACAGCATCTGCGCGCGCGCGCGCGGAGGCGCGGCTCGTCAACC 300
Qy 359 TCAGGAGCAGGAGTGTCTTCCCGCAGCTGGGAGGCTTATTCTCGGGCTCCGAGGGCAGC 418
Db 301 TCAGGAGCAGGAGTGTCTTCCCGCAGCTGGGAGGCTTATTCTCGGGCTCCGAGGGCAGC 360
Qy 419 TGAAGCCCGGGGCGCGCATCTTCTTTTCGAGCGGAGAGATGTCTTCGAGCACCCTCCGCT 478
Db 361 TGAAGCCCGGGGCGCGCATCTTCTTTTCGAGCGGAGAGATGTCTTCGAGCACCCTCCGCT 420
Qy 479 GGGCCCGGAGAGCGGAGTGTGCGAGCGGCTTCGAGCACCCTCCGCGCGCGCGCAGCT 538
Db 421 GGGCCCGGAGAGCGGAGTGTGCGAGCGGCTTCGAGCACCCTCCGCGCGCGCGCAGCT 480
Qy 539 ACTGCGAGAGCTGTGGCGGAGCGGAGGCGCGCGCGCGCGCGGAGGCTGTCTGCTGTGG 598

Db 481 ACTGCAGACGCTGCGGACGAGGCCCCCGGCGCCACCGGGCAGGCGTCTGTCGCTGCTGG 540
QY 599 CGGGCAGGCTGCTGGAGCAGGAGCGCGGAGTTCGCCACGCTTCGTGGTGGCTCTGCA 658
Db 541 CGGGCAGGCTGCTGGAGCAGGAGCGCGGAGTTCGCCACGCTTCGTGGTGGCTCTGCA 600
QY 659 TCGAGAACGCTCATGACCTCTTCTCCAAG 690
Db 601 TCGAGAACGCTCATGACCTCTTCTCCAAG 632

RESULT 4
US-09-880-107-713/c
; Sequence 713, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene-Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR FILING DATE: 2000-06-14
; PRIOR FILING DATE: 2000-06-14
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 713
; LENGTH: 251
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA292773
US-09-880-107-713

Query Match 4.2%; Score 35; DB 10; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 795 ATGTAATCTCTCAAGAAATAAAGGAAGCCCAAGAG 829
Db 41 ATGTAATCTCTCAAGAAATAAAGGAAGCCCAAGAG 7

RESULT 5
US-09-880-107-2178
; Sequence 2178, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene-Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR FILING DATE: 2000-06-14
; PRIOR FILING DATE: 2000-06-14
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2178
; LENGTH: 3394
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L22548
US-09-880-107-2178

Query Match 4.2%; Score 35; DB 10; Length 3394;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 795 ATGTAATCTCTCAAGAAATAAAGGAAGCCCAAGAG 829
Db 2187 ATGTAATCTCTCAAGAAATAAAGGAAGCCCAAGAG 2221

RESULT 6
US-09-918-995-23271
; Sequence 23271, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR FILING DATE: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23271
; LENGTH: 297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(297)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-23271

Query Match 3.5%; Score 29; DB 11; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 795 ATGTAATCTCTCAAGAAATAAAGGAAGCC 823
Db 241 ATGTAATCTCTCAAGAAATAAAGGAAGCC 269

RESULT 7
US-10-060-036-4
; Sequence 4, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: n = A,T,C or G
US-10-060-036-4

Query Match 3.3%; Score 27; DB 14; Length 434;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTCGCGGCGAGATGACATCTCGGCC 27

Db 210 CCTGGCGGCAGATGACATCTGGCC 236
|||||

RESULT 8

US-10-060-036-64

; Sequence 64, Application US/10060036

; Publication No. US20030073144A1

; GENERAL INFORMATION:

; APPLICANT: Benson, Darin R.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Lodes, Michael J.

; APPLICANT: Persing, David H.

; APPLICANT: Hepler, William T.

; APPLICANT: Jiang, Yuqiu

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.566

; CURRENT APPLICATION NUMBER: US/10/060,036

; CURRENT FILING DATE: 2002-01-30

; NUMBER OF SEQ ID NOS: 4560

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 64

; LENGTH: 574

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-060-036-64

Query Match 3.3%; Score 27; DB 14; Length 574;

Best Local Similarity 100.0%; Pred. No. 0.0016;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGGCGGCAGATGACATCTGGCC 27

Db 209 CCTGGCGGCAGATGACATCTGGCC 235

RESULT 9

US-10-060-036-144

; Sequence 144, Application US/10060036

; Publication No. US20030073144A1

; GENERAL INFORMATION:

; APPLICANT: Benson, Darin R.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Lodes, Michael J.

; APPLICANT: Persing, David H.

; APPLICANT: Hepler, William T.

; APPLICANT: Jiang, Yuqiu

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.566

; CURRENT APPLICATION NUMBER: US/10/060,036

; CURRENT FILING DATE: 2002-01-30

; NUMBER OF SEQ ID NOS: 4560

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 144

; LENGTH: 4551

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-060-036-144

Query Match 3.3%; Score 27; DB 14; Length 4551;

Best Local Similarity 100.0%; Pred. No. 0.0012;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGGCGGCAGATGACATCTGGCC 27

Db 3874 CCTGGCGGCAGATGACATCTGGCC 3900

RESULT 10

US-10-042-347-6

; Sequence 6, Application US/10042347

; Publication No. US20030114370A1

; GENERAL INFORMATION:

; APPLICANT: O'Reilly, Michael S.

; APPLICANT: Folkman, M. Judah

; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Endostatin Protein and Peptide Fra

; FILE REFERENCE: 05213-0880 (43170-249874)

; CURRENT APPLICATION NUMBER: US/10/042,347

; CURRENT FILING DATE: 2002-01-11

; PRIOR APPLICATION NUMBER: US 09/315,689

; PRIOR FILING DATE: 1999-05-20

; PRIOR APPLICATION NUMBER: US 60/106,343

; PRIOR FILING DATE: 1998-10-30

; PRIOR APPLICATION NUMBER: US 09/154,302

; PRIOR FILING DATE: 1998-09-16

; PRIOR APPLICATION NUMBER: US 08/740,168

; PRIOR FILING DATE: 1996-10-22

; PRIOR APPLICATION NUMBER: US 60/005,835

; PRIOR FILING DATE: 1995-10-23

; PRIOR APPLICATION NUMBER: US 60/023,070

; PRIOR FILING DATE: 1996-08-02

; PRIOR APPLICATION NUMBER: US 60/026,263

; PRIOR FILING DATE: 1996-09-17

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 534

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-042-347-6

Query Match 3.1%; Score 26; DB 14; Length 534;

Best Local Similarity 100.0%; Pred. No. 0.005;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 226 GACTTCCAGTGCTTCCAGCAGCGCG 251

Db 76 GACTTCCAGTGCTTCCAGCAGCGCG 101

RESULT 11

US-10-131-241-59

; Sequence 59, Application US/10131241

; Publication No. US20030012792A1

; GENERAL INFORMATION:

; APPLICANT: Holaday, John W.

; APPLICANT: Fortier, Anne H.

; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer

; FILE REFERENCE: 05213-0344 43170-271565

; CURRENT APPLICATION NUMBER: US/10/131,241

; CURRENT FILING DATE: 2002-07-22

; PRIOR APPLICATION NUMBER: US 09/413,049

; PRIOR FILING DATE: 1999-10-06

; PRIOR APPLICATION NUMBER: US 09/316,802

; PRIOR FILING DATE: 1999-05-21

; PRIOR APPLICATION NUMBER: US 60/086,586

; PRIOR FILING DATE: 1998-05-22

; NUMBER OF SEQ ID NOS: 65

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 59

; LENGTH: 537

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-131-241-59

Query Match 3.1%; Score 26; DB 14; Length 537;

Best Local Similarity 100.0%; Pred. No. 0.005;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 226 GACTTCCAGTGCTTCCAGCAGCGCG 251

Db 76 GACTTCCAGTGCTTCCAGCAGCGCG 101

RESULT 12
US-10-042-347-4
; Sequence 4, Application US/10042347
; Publication No. US20030114370A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Endostatin Protein and Peptide Fragments
; FILE REFERENCE: 05213-0880 (43170-249874)
; CURRENT APPLICATION NUMBER: US/10/042,347
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: US 09/315,689
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 60/106,343
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 09/154,302
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 08/740,168
; PRIOR FILING DATE: 1996-10-22
; PRIOR APPLICATION NUMBER: US 60/005,835
; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-347-4

Query Match 3.1%; Score 26; DB 14; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.005;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 226 GACTTCCAGTGCTTCCAGCAGGCGG 251
DB 88 GACTTCCAGTGCTTCCAGCAGGCGG 113

RESULT 13
US-10-292-418-3
; Sequence 3, Application US/10292418
; Publication No. US20030139365A1
; GENERAL INFORMATION:
; APPLICANT: Lo, Kin-Ming
; APPLICANT: Li, Yue
; APPLICANT: Gillies, Stephen D
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as Peptides
; FILE REFERENCE: LEX-006C1
; CURRENT APPLICATION NUMBER: US/10/292,418
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/383,315
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: US 60/097,883
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(549)
; OTHER INFORMATION: endostatin
US-10-292-418-3

Query Match 3.1%; Score 26; DB 12; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.005;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 226 GACTTCCAGTGCTTCCAGCAGGCGG 251
DB 88 GACTTCCAGTGCTTCCAGCAGGCGG 113

RESULT 14
US-10-131-241-53
; Sequence 53, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-241-53

Query Match 3.1%; Score 26; DB 14; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.005;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 226 GACTTCCAGTGCTTCCAGCAGGCGG 251
DB 88 GACTTCCAGTGCTTCCAGCAGGCGG 113

RESULT 15
US-10-080-797-2
; Sequence 2, Application US/10080797
; Publication No. US20020183253A1
; GENERAL INFORMATION:
; APPLICANT: Campochiaro, Peter A.
; APPLICANT: Dixon, Katharine H.
; APPLICANT: Brazzell, Romulus K.
; TITLE OF INVENTION: METHOD FOR TREATING OCULAR DISEASES
; FILE REFERENCE: 4-31881A
; CURRENT APPLICATION NUMBER: US/10/080,797
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Human
US-10-080-797-2

Query Match 3.1%; Score 26; DB 13; Length 551;
Best Local Similarity 100.0%; Pred. No. 0.005;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 226 GACTTCCAGTGCTTCCAGCAGGCGG 251
DB 87 GACTTCCAGTGCTTCCAGCAGGCGG 112

Mon Aug 18 10:26:55 2003

us-09-938-391-1.oli.rnpb

Page 6

Search completed: August 17, 2003, 19:58:08
Job time : 320.875 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 16:25:28 ; Search time 84.4574 Seconds
(without alignments)
4332.442 Million cell updates/sec

Title: US-09-938-391-1
Perfect score: 829
Sequence: 1 cccgtggcggcagatgacat.....aataaaggagccaaagag 829

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	4.2	3394	1	US-08-159-784-4
2	35	4.2	4031	1	US-08-159-784-1
3	26	3.1	534	4	US-09-315-689-6
4	26	3.1	546	4	US-09-315-689-4
5	26	3.1	552	3	US-09-206-059-30
6	19	2.3	801	2	US-08-770-379-16
7	19	2.3	801	3	US-08-757-669A-16
8	19	2.3	801	4	US-09-298-568-3
9	19	2.3	801	4	US-09-230-371A-16
10	19	2.3	3500	3	US-09-197-636-7
11	19	2.3	4803	3	US-09-197-636-1
12	19	2.3	4803	3	US-09-197-636-3
13	19	2.3	112132	4	US-09-741-150-3
14	18	2.2	273	4	US-09-313-294A-2718
15	18	2.2	288	3	US-08-956-307B-10
16	18	2.2	292	3	US-08-956-307B-9
17	18	2.2	558	3	US-09-449-293-3
18	18	2.2	558	4	US-09-775-325-3
19	18	2.2	565	3	US-08-985-526-37
20	18	2.2	573	4	US-09-561-500-12
21	18	2.2	573	4	US-09-561-108-12
22	18	2.2	573	4	US-09-561-526-12
23	18	2.2	573	4	US-09-561-499-12
24	18	2.2	774	3	US-08-956-307B-12
25	18	2.2	778	3	US-08-956-307B-11
26	18	2.2	1557	4	US-09-252-991A-16492
27	18	2.2	1785	4	US-09-252-991A-16007, A

c 28	18	2.2	2028	4	US-09-252-991A-16110	Sequence 16110, A
c 29	18	2.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c 30	18	2.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c 31	18	2.2	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c 32	18	2.2	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c 33	17	2.1	522	4	US-09-252-991A-16260	Sequence 16260, A
c 34	17	2.1	630	4	US-09-252-991A-5585	Sequence 5585, Ap
c 35	17	2.1	690	4	US-09-252-991A-2533	Sequence 2533, Ap
c 36	17	2.1	753	4	US-09-252-991A-5613	Sequence 5613, Ap
c 37	17	2.1	870	4	US-09-252-991A-2385	Sequence 2385, Ap
c 38	17	2.1	918	4	US-09-252-991A-5525	Sequence 5525, Ap
c 39	17	2.1	951	4	US-09-252-991A-16122	Sequence 16122, A
c 40	17	2.1	1023	4	US-09-252-991A-14229	Sequence 14229, A
c 41	17	2.1	1065	4	US-09-252-991A-12208	Sequence 12208, A
c 42	17	2.1	1111	4	US-09-620-312D-797	Sequence 797, App
c 43	17	2.1	1233	4	US-09-252-991A-14035	Sequence 14035, A
c 44	17	2.1	1326	4	US-09-252-991A-2299	Sequence 2299, Ap
c 45	17	2.1	1356	4	US-09-252-991A-14188	Sequence 14188, A

ALIGNMENTS

RESULT 1
US-08-159-784-4
; Sequence 4, Application US/08159784
; Patent No. 5643783
; GENERAL INFORMATION:
; APPLICANT: Bjorn R. Olsen
; TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,784
; FILING DATE: December 1, 1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: John F. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00246/170001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3394
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-159-784-4

Query Match 4.2%; Score 35; DB 1; Length 3394;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCTCTCAAGAAATAAAAGGAGCCCAAGAG 829
|||||
Db 2187 ATGTAATCTCTCAAGAAATAAAAGGAGCCCAAGAG 2221
|||||

RESULT 2
US-08-159-784-1
; Sequence 1, Application US/08159784
; Patent No. 5643783
; GENERAL INFORMATION:
; APPLICANT: Bjorn R. Olsen
; TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: Wordperfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,784
; FILING DATE: December 1, 1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: John F. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00246/170001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4031
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-159-784-1

Query Match 4.2%; Score 35; DB 1; Length 4031;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCCTCAAGAAATAAAGGAGCCAAAGAG 829
Db 3987 ATGTAATCCTCAAGAAATAAAGGAGCCAAAGAG 4021

RESULT 3
US-09-315-689-6
; Sequence 6, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Folkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-315-689-6

Query Match 3.1%; Score 26; DB 4; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 226 GACTTCCAGTGCTTCCAGCAGCGCG 251
Db 76 GACTTCCAGTGCTTCCAGCAGCGCG 101

RESULT 4
US-09-315-689-4
; Sequence 4, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Folkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-315-689-4

Query Match 3.1%; Score 26; DB 4; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 226 GACTTCCAGTGCTTCCAGCAGCGCG 251
Db 88 GACTTCCAGTGCTTCCAGCAGCGCG 113

RESULT 5
US-09-206-059-30
; Sequence 30, Application US/09206059
; Patent No. 6201104
; GENERAL INFORMATION:
; APPLICANT: MacDonald, Nicholas
; APPLICANT: Sim, Kim Lee
; TITLE OF INVENTION: Angiogenesis-Inhibiting Protein Binding Peptides and
; FILE REFERENCE: 05213-0370
; CURRENT APPLICATION NUMBER: US/09/206,059
; CURRENT FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-206-059-30

Query Match 3.1%; Score 26; DB 3; Length 552;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 226 GACTTCCAGTGCTTCCAGCAGCGCG 251
Db 88 GACTTCCAGTGCTTCCAGCAGCGCG 113

RESULT 6
US-08-770-379-16/c
; Sequence 16, Application US/08770379
; Patent No. 5849564
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.

; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,371A
; PRIOR FILING DATE: 1999-11-17
; PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 16
; LENGTH: 801
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-16

Query Match 2.3%; Score 19; DB 4; Length 801;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 716 CGGGGAGGGGGCGCCGC 734
DB 484 CGGGGAGGGGGCGCCGC 466

RESULT 10

US-09-197-636-7/c
; Sequence 7, Application US/09197636
; Patent No. 6239267
; GENERAL INFORMATION:
; APPLICANT: DUCKWORTH, DAVID
; APPLICANT: HAYES, PHILIP
; APPLICANT: MEADOWS, HELEN
; APPLICANT: DAVIS, JOHN
; TITLE OF INVENTION: NOVEL COMPOUNDS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: US
; ZIP: 19482-0980
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/197,636
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 9805137.8
; FILING DATE: 12-MAR-1998
; APPLICATION NUMBER: UK 9815791.0
; FILING DATE: 21-JUL-1998
; APPLICATION NUMBER: UK 9819278.4
; FILING DATE: 03-SEP-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GP-30075
; TELEPHONE: 601-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169

; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3500 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: cdna
US-09-197-636-7

Query Match 2.3%; Score 19; DB 3; Length 3500;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 CCACACCCACCCACACC 144
DB 324 CCACACCCACCCACACC 306

RESULT 11

US-09-197-636-1/c
; Sequence 1, Application US/09197636
; Patent No. 6239267
; GENERAL INFORMATION:
; APPLICANT: DUCKWORTH, DAVID
; APPLICANT: HAYES, PHILIP
; APPLICANT: MEADOWS, HELEN
; APPLICANT: DAVIS, JOHN
; TITLE OF INVENTION: NOVEL COMPOUNDS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: US
; ZIP: 19482-0980
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/197,636
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 9805137.8
; FILING DATE: 12-MAR-1998
; APPLICATION NUMBER: UK 9815791.0
; FILING DATE: 21-JUL-1998
; APPLICATION NUMBER: UK 9819278.4
; FILING DATE: 03-SEP-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GP-30075
; TELEPHONE: 601-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4803 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-197-636-1

Query Match 2.3%; Score 19; DB 3; Length 4803;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 CCACACCCACCCACACC 144
DB 324 CCACACCCACCCACACC 306

RESULT 12

```
US-09-197-636-3/c
; Sequence 3, Application US/09197636
; Patent No. 6239267
; GENERAL INFORMATION:
; APPLICANT: DUCKWORTH, DAVID
; APPLICANT: HAYES, PHILIP
; APPLICANT: MEADOWS, HELEN
; APPLICANT: DAVIS, JOHN
; TITLE OF INVENTION: NOVEL COMPOUNDS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: US
; ZIP: 19482-0980
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/197,636
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 9805137.8
; FILING DATE: 12-MAR-1998
; APPLICATION NUMBER: UK 9815791.0
; FILING DATE: 21-JUL-1998
; APPLICATION NUMBER: UK 9819278.4
; FILING DATE: 03-SEP-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GP-30075
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 601-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4803 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-197-636-3

Query Match 2.3%; Score 19; DB 3; Length 4803;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 CCACACCCACCCACACC 144
DB 324 CCACACCCACCCACACC 306

RESULT 13
US-09-741-150-3
; Sequence 3, Application US/09741150
; Patent No. 6436689
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000968
; CURRENT APPLICATION NUMBER: US/09/741,150
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0

us-09-938-391-1.oli.rni

; SEQ ID NO 3
; LENGTH: 112132
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(112132)
; OTHER INFORMATION: n = A,T,C or G
US-09-741-150-3

Query Match 2.3%; Score 19; DB 4; Length 112132;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 CACCCACCCACCCACCCAC 147
DB 32958 CACCCACCCACCCACCCAC 32976

RESULT 14
US-09-313-294A-2718
; Sequence 2718, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 2718
; LENGTH: 273
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700553162H1
; NAME/KEY: unsure
; LOCATION: 88, 189, 238, 243, 246, 255, 260
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-2718

Query Match 2.2%; Score 18; DB 4; Length 273;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 566 CGCGCGCCACCGGCAGG 583
DB 98 CGCGCGCCACCGGCAGG 115

RESULT 15
US-08-956-307B-10
; Sequence 10, Application US/08956307B
; Patent No. 6090911
; GENERAL INFORMATION:
; APPLICANT: Petka, Wendy A.
; APPLICANT: Tirrell, David A.
; APPLICANT: Kevin P. McGrath
; TITLE OF INVENTION: REVERSIBLE HYDROGELS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```

```

;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,307B
; FILING DATE: 22-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07880/033001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-956-307B-10

Query Match      2.2%; Score 18; DB 3; Length 288;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      265 GCGGGCACCTTCGGGCC 282
Db      54 GCGGGCACCTTCGGGCC 71

```

Search completed: August 17, 2003, 19:49:19
Job time : 94.4574 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 13:17:18 ; Search time 4162.37 Seconds
(without alignments)
8147.780 Million cell updates/sec

Title: US-09-938-391-1

Perfect score: 829

Sequence: 1 cctctggcggcgatgacat.....aataaaggaagcacaagag 829

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.in.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.man.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	829	100.0	829	6	AX399629	AX399629 Sequence
2	555	66.9	555	6	AX399631	AX399631 Sequence
3	504.8	60.9	2637	9	BC033715	BC033715 Homo sapi
4	503.2	60.7	5408	9	AF018082	AF018082 Homo sapi
5	503.2	60.7	5929	9	AF018081	AF018081 Homo sapi
6	501.6	60.5	2364	9	AK098216	AK098216 Homo sapi
7	496.8	59.9	3394	6	AX409531	AX409531 Sequence
8	496.8	59.9	3394	6	IS1045	IS1045 Sequence 4
9	496.8	59.9	3394	9	HUMCOL18AX	L22548 Human colla
10	493.2	59.5	900	6	AX370851	AX370851 Sequence
11	436.8	52.7	1615	10	MUSCOL3714	U03714 Mus musculu
12	436.8	52.7	4437	10	MMUSCOLLAG	D17546 Mus musculu
13	436.8	52.7	4802	10	BC043697	BC043697 Mus muscu
14	423	51.0	4031	6	IS1044	IS1044 Sequence 1
15	423	51.0	4031	10	MUSCOLPRO	L22545 Mouse alpha
16	420	50.7	564	9	AF416592	AF416592 Homo sapi
17	418	50.4	546	6	AR193165	AR193165 Sequence
18	418	50.4	549	6	AX100086	AX100086 Sequence
19	418	50.4	552	6	AX395662	AX395662 Sequence
20	418	50.4	552	6	AX473835	AX473835 Sequence
21	418	50.4	552	6	BD081407	BD081407 Fused pro
22	418	50.4	786	9	AF282883	AF282883 Homo sapi
23	411.6	49.7	555	9	AF184060	AF184060 Homo sapi
24	410.8	49.6	534	6	AR193166	AR193166 Sequence
25	410.8	49.6	537	6	AX100092	AX100092 Sequence
26	409.4	49.4	851	10	AF189709	AF189709 Rattus no
27	398.4	48.1	650	6	AX370853	AX370853 Sequence
28	377.8	45.6	624	6	AX084406	AX084406 Sequence
29	376.8	45.5	573	6	AR183471	AR183471 Sequence
30	376.8	45.5	573	6	AR183520	AR183520 Sequence
31	376.8	45.5	573	6	AR217279	AR217279 Sequence
32	376.8	45.5	573	6	AR282751	AR282751 Sequence
33	376.8	45.5	573	6	AX042272	AX042272 Sequence
34	376.8	45.5	573	6	AX490666	AX490666 Sequence
35	376.6	45.4	565	6	AR100642	AR100642 Sequence
36	376.6	45.4	565	6	E34073	E34073 Carrier/DNA
37	376.2	45.4	558	6	AR268609	AR268609 Sequence
38	376.2	45.4	624	6	AX128409	AX128409 Sequence
39	375.8	45.3	555	6	BD102800	BD102800 Process f
40	375.8	45.3	555	6	BD128313	BD128313 Endostat i
41	373.4	45.0	552	10	AF257775	AF257775 Mus muscu
42	350.8	42.3	515	10	RNO236873	AJ236873 Rattus no
43	349.8	42.2	5279	5	AF083440	AF083440 Gallus ga
44	251	30.3	5598	5	AB047066	AB047066 Xenopus l
45	248.8	30.0	2037	5	DRE494837	AJ494837 Danio rer

ALIGNMENTS

RESULT 1	AX399629	Sequence 1 from Patent EP1191036.	829 bp	DNA	linear	PAT 06-JUN-2002
LOCUS	AX399629	Sequence 1 from Patent EP1191036.				
DEFINITION	AX399629	Sequence 1 from Patent EP1191036.				
ACCESSION	AX399629	Sequence 1 from Patent EP1191036.				
VERSION	AX399629.1	GI:21335409				
KEYWORDS		Canis familiaris (dog)				
SOURCE		Canis familiaris				
ORGANISM		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
REFERENCE		Mammalia; Euthera; Carnivora; Fissipedia; Canidae; Canis.				
AUTHORS		Sheppard, M.G. and Tong, X.				
TITLE		Methods and compositions for diagnosing and treating disorders involving angiogenesis				

JOURNAL Patent: EP 1191036-A 1 27-MAR-2002;
Pfizer Products Inc. (US)
FEATURES Location/Qualifiers
source 1..829
/organism="Canis familiaris"
/mol_type="genomic DNA"
/db_xref="taxon:9615"
BASE COUNT 124 a 314 c 278 g 113 t
ORIGIN

Query Match 100.0%; Score 829; DB 6; Length 829;
Best Local Similarity 100.0%; Pred. No. 7.8e-100;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTGGCGGCGAGATGACATCTCTGGCGGCGCCCGCGCTGCTGACCCCGACCGCTAC 60
DB 1 CCCTGGCGGCGAGATGACATCTCTGGCGGCGCCCGCGCTGCTGACCCCGACCGCTAC 60
QY 61 CCGGGGCGCGCAGCAGCGCTCTAGTGCACTTCCAGCGGCTCGCCCACTGGTGGG 120
DB 61 CCGGGGCGCGCAGCAGCGCTCTAGTGCACTTCCAGCGGCTCGCCCACTGGTGGG 120
QY 121 CCGGTCCACACCCACACCCACACCCAGGACTTCCAGTGTGTGCACTGTGTGGCC 180
DB 121 CCGGTCCACACCCACACCCACACCCAGGACTTCCAGTGTGTGCACTGTGTGGCC 180
QY 181 CTGAACAGCCCGCAGCGGCGGCGATCGAGGATCCGGGAGCGGACTTCCAGTGTTC 240
DB 181 CTGAACAGCCCGCAGCGGCGGCGATCGAGGATCCGGGAGCGGACTTCCAGTGTTC 240
QY 241 CAGCAGCGCGCGCGCGGCGCTTCCGCGCACTTCCGCGCTTCCGCGCTG 300
DB 241 CAGCAGCGCGCGCGCGGCGCTTCCGCGCACTTCCGCGCTTCCGCGCTG 300
QY 301 CAGCAGCTTACAGATCGTGTGCGCGCGCGACCGCAACCGGGGTGCCGCTCAACCTC 360
DB 301 CAGCAGCTTACAGATCGTGTGCGCGCGCGACCGCAACCGGGGTGCCGCTCAACCTC 360
QY 361 AGGACGAGGTGTCTTCCCGAGCTGGAGGCGCTTATTCGCGGCTCCGAGGCGAGCTG 420
DB 361 AGGACGAGGTGTCTTCCCGAGCTGGAGGCGCTTATTCGCGGCTCCGAGGCGAGCTG 420
QY 421 AAGCCCGGGCGCGCATCTTCTCTTTCAGCGGAGAGATGCTCTGAGCAACCCCGCTGG 480
DB 421 AAGCCCGGGCGCGCATCTTCTCTTTCAGCGGAGAGATGCTCTGAGCAACCCCGCTGG 480
QY 481 CCGCGGAAGAGCGTGTGGCAAGCGTCCGACCCAGCGGGCGCGCTGACCGAGCTAC 540
DB 481 CCGCGGAAGAGCGTGTGGCAAGCGTCCGACCCAGCGGGCGCGCTGACCGAGCTAC 540
QY 541 TGGAGACGTGGCGGAGGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
DB 541 TGGAGACGTGGCGGAGGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
QY 601 GGCAGGCTGTGGAGCAGGAGGCGCGGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGG 660
DB 601 GGCAGGCTGTGGAGCAGGAGGCGCGGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGG 660
QY 661 GAGAACAGCGTCATGACTCTCTTCCAAAGTAGGGCGCGGCGGCGGCGGCGGCGGCGG 720
DB 661 GAGAACAGCGTCATGACTCTCTTCCAAAGTAGGGCGCGGCGGCGGCGGCGGCGGCGG 720
QY 721 GAGGGGCGCGCGGAGGAGATCCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
DB 721 GAGGGGCGCGCGGAGGAGATCCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
QY 781 CACCGTCACGTTTAAATGTAATCTCTCAAGAAATAAAGGAAGCAAGAG 829
DB 781 CACCGTCACGTTTAAATGTAATCTCTCAAGAAATAAAGGAAGCAAGAG 829

RESULT 2
AX399631

LOCUS AX399631 555 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 3 from Patent EP1191036.
ACCESSION AX399631
VERSION AX399631.1 GI:21335410
KEYWORDS Canis familiaris (dog)
SOURCE Canis familiaris
ORGANISM Canis familiaris
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE 1
AUTHORS Shepard, M.G. and Tong, X.
TITLE Methods and compositions for diagnosing and treating disorders involving angiogenesis
JOURNAL Patent: EP 1191036-A 3 27-MAR-2002;
Pfizer Products Inc. (US)
FEATURES Location/Qualifiers
source 1..555
/organism="Canis familiaris"
/mol_type="genomic DNA"
/db_xref="taxon:9615"
BASE COUNT 78 a 203 c 191 g 83 t
ORIGIN

Query Match 66.9%; Score 555; DB 6; Length 555;
Best Local Similarity 100.0%; Pred. No. 8.4e-64;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 CACACCCACGAGACTTCCAGCTGTGTGCACTGTGTGGCGCTGCAACAGCCGCGAGCG 198
DB 1 CACACCCACGAGACTTCCAGCTGTGTGCACTGTGTGGCGCTGCAACAGCCGCGAGCG 60
QY 199 GCGCGCATCCGAGGATCCGCGGAGCGGACTTCCAGTGTTCAGAGCGCGCGCGCG 258
DB 61 GCGCGCATCCGAGGATCCGCGGAGCGGACTTCCAGTGTTCAGAGCGCGCGCGCG 120
QY 259 GCGCTGCGCGCACTTCCGCGCTTCTGTGTGTGCTGCGGCTCGAGACCTTACAGCATC 318
DB 121 GCGCTGCGCGCACTTCCGCGCTTCTGTGTGTGCTGCGGCTCGAGACCTTACAGCATC 180
QY 319 GTGCGCGCGCGCGACCGCGCGGCTGCGGCTGCTCAACCTCAGGACGAGGTGCTTTC 378
DB 181 GTGCGCGCGCGCGACCGCGCGGCTGCGGCTGCTCAACCTCAGGACGAGGTGCTTTC 240
QY 379 CCGAGCTGGAGGCGCTTATTCCTCGGCTCCGAGGCGGAGCTGAAGCCCGGGCGCGCATC 438
DB 241 CCGAGCTGGAGGCGCTTATTCCTCGGCTCCGAGGCGGAGCTGAAGCCCGGGCGCGCATC 300
QY 439 TTCTCTTTCGAGCGGAGAGATGCTGTGAGCAACCCCGCTGCGGCCCGGAGAGCGTGTGG 498
DB 301 TTCTCTTTCGAGCGGAGAGATGCTGTGAGCAACCCCGCTGCGGCCCGGAGAGCGTGTGG 360
QY 499 CAGGCTCCGACCCCGAGCGGCGGCTGACCGGAGCTACTGCGAGAGCTGGCGGAGCG 558
DB 361 CAGGCTCCGACCCCGAGCGGCGGCTGACCGGAGCTACTGCGAGAGCTGGCGGAGCG 420
QY 559 GAGGCGCGCGCGCGCACCGCGGAGGCGTGTGTGCTGCGGCGGAGGCTGCTGGAGCAG 618
DB 421 GAGGCGCGCGCGCGCACCGCGGAGGCGTGTGTGCTGCGGCGGAGGCTGCTGGAGCAG 480
QY 619 GAGGCGCGGAGCTGCGCGCACCGCTTCTGTGTGTGCTTGTGATCGAGAACAGCGTCAAGCC 678
DB 481 GAGGCGCGGAGCTGCGCGCACCGCTTCTGTGTGTGCTTGTGATCGAGAACAGCGTCAAGCC 540
QY 679 TCCTTCTCCAAGTAG 693
DB 541 TCCTTCTCCAAGTAG 555

RESULT 3
BC033715 2637 bp mRNA linear PRI 08-JUL-2002
LOCUS BC033715
DEFINITION Homo sapiens, Similar to collagen, type XVIII, alpha 1, clone IMAGE:4425380, mRNA, partial cds.

	Query Match	60.7%; Score 503.2; DB 9; Length 5408;	SOURCE	Homo sapiens (human)
	Best Local Similarity	77.6%; Pred. No. 2.7e-57;	ORGANISM	Eukaryotes
	Matches 664; Conservative	0; Mismatches 153; Indels 39; Gaps 3;		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Qy	1	CCTGGCGGCAGATCACATCTGTCGCCGCCCCCGCCTGTGGACCCCAGCCCTTAC 60 	REFERENCE	1 (bases 1 to 5929) <i>Saaresla,J., Viikarppa,R., Rehn,M., Purmonen,S. and Pihlajaniemi,T. Complete primary structure of two variant forms of human type XVIII collagen and tissue-specific differences in the expression of the corresponding transcripts</i>
Dd	3393	CCTGGCGGCAGATCACATCTGTCGCCGCCCCCGCCTGTGGACCCCAGCCCTTAC 3452 	TITLE	
Qy	61	CCCCGGGCCCCGACCACCGGCTCTACTGFGCACTTTCCAGCCGGCTCGCCCACTGGTGCG 120 	JOURNAL	Matrix Biol. 16 (6), 319-328 (1998)
Dd	3453	CCCGAGCCCCGACACACAGCTCTTACTGTCACCTCGCGCCGGCGACCAAG---- 3508 	MEDLINE	98164096
			PUBMED	9503365
Qy	121	CCCGTCCACCCACACACCCACACCCACAGGAATTCCAGCTGGTGCTGCACCTGGTGGCC 180 	REFERENCE	2 (bases 1 to 5929) <i>Saaresla,J., Viikarppa,R., Rehn,M., Purmonen,S. and Pihlajaniemi,T. Direct Submission</i>
Dd	3509	-----CCACCCGCCCCACAGCACCCGCGAATTCCAGCCGGTGTCTCCACTGGTTGCG 3560 	AUTHORS	
			TITLE	

QY	361	AGGACGAGGTGCTTCCACGCTGGAGGCGCTTATTTCTCGGCTCCGAGGCCAGCTG	420		
Db	4262	AAGGACGAGCTGCTGTTTCCAGCTGGAGGCTCTGTTTCTCAGGCTCTGAGGCTCGCGCTG <td>4321</td>	4321		
QY	421	AAGCCCGGGCCCGCATCTTCTTTTCGACGGCAGAGATGCTCTCGACGACCCCGCCTGG <td>480</td>	480		
Db	4322	AAGCCCGGGCCACGCACTTCTCTCTTTGACGGCAAGACGCTCTCTGAGGCAACCCACCTGG <td>4381</td>	4381		
QY	481	CCCGGAAGAGCTGTGGCAGCGCTCCGACCCCGAGCGGGCGCGCTTGACCGACAGCTAC <td>540</td>	540		
Db	4382	CCCCAGAAGAGCTGTGGCATGAGCTTGGACCCCAACGGGCGCAGGCTGACCGAGAGCTAC <td>4441</td>	4441		
QY	541	TGCGAGACGTGGCGGACGAGGCGCCCGCGGCCACCGGGCAGGCGTCTGCTGCTGCGCG <td>600</td>	600		
Db	4442	TGTGACAGCTGGCGGACGAGGCTCCTCTGGCCACCGGCCAGGCTCTCTGCTGCTGGG <td>4501</td>	4501		
QY	601	GGCAGCTGCTGAGCAGAGGCGCGGAGCTGCGGCCACCGCTTCGTGTGCTCTGTCATC <td>660</td>	660		
Db	4502	GGCAGCTCTCTGGGGCAGAGTGC CGGAGTGCATCAGCCATCAGCTGCTGCTGCTGCTG <td>4551</td>	4551		
QY	661	GAGAACAGCTCATGACCTCTTCTCAAGTAGGCGCGCGCGCCACCGACAGGCGGGG <td>720</td>	720		
Db	4562	GAGAACAGCTTCATGACTGCTTCAAGTAGCCACCGCTCGATGCGGATGCGCGAGAGG <td>4621</td>	4621		
QY	721	GAGGGGCGCGCCGACGAGCATCGCGCCCGCGGG----- <td>756</td>	756		
Db	4622	ACCGCGGCTCGAGGAAGCCCGCCACCGTGGGCGAGGAGCGCGCGCCCGCTGCGCC <td>4681</td>	4681		
QY	757	--GGGCTGCGCGGACGCTTGGCTGCGACCG--TCAGCTTTAATGTAATCTCAAGAAATA <td>813</td>	813		
Db	4682	CAGGACCTGGCTGCCATCTTCTCTATAGTTTCACTTTTCATGTAATCTCTCAAGAAATA <td>4741</td>	4741		
QY	814	AAAGGAAGCCAAAGAG <td>829</td>	829		
Db	4742	AAAGGAAGCCAAAGAG <td>4757</td>	4757		
RESULT 6					
AK098216					
LOCUS					
DEFINITION	Homo sapiens cDNA FLJ40897 fis, clone UTERU002964, highly similar to Homo sapiens type XVIII collagen (COL18A1) mRNA.	linear	PRI 15-JUL-2002		
ACCESSION	AK098216				
VERSION	AK098216.1				
KEYWORDS	oligo capping; fis (full insert sequence).				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Mammalia; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
JOURNAL					
TITLE					
COMMENT					

Db 3532 CTGAAGGACGAGGTGCTAICTCCAGCTGGGACTCCCTGTGTTTTCTGGCTCCCAGGGTCAA 3591
QY 418 CTGAAGCCCGGGCCCGCATCTTCTTTTCGAGCGAGAGATGCTCCTGACGACACCCCGCC 477
Db 3592 GTGCAACCCCGGGCCCGCATCTTCTTTTGACGGCAGAGATGCTCTGACACACCCAGCC 3651
QY 478 TGGCCCCGGAAGAGCGTGTGGCACGGCTCCGACCCCAAGCGGGCGCCGCTGACCCGACAGC 537
Db 3652 TGGCCGACAGAGAGCGTATGGCACGGCTCGGACCCCAAGTGGCGGAGGCTGATGGAGAGT 3711
QY 538 TACTGCGAGACGTGGCGGACGAGGCCCGCGCGGCCACCGGGCAGGGTCTGCTGCTG 597
Db 3712 TACTGTGAGACATGGCGAACTGAAACTACTTGGGGCTACAGGTCAAGGCTCCTCCCTGCTG 3771
QY 598 GCGGGCAGGCTGTGGAGCAGGAGCGCGAGCTGCCGCCACGCCCTTCGTGGTGTCTGC 657
Db 3772 TCAGGCAGGCTCTGGAACAGAAAGCTGCGAGTGCACACAGCTACATCGTCTGTGC 3831
QY 658 ATCGAAGACGCTCATGACCTCTTCTTCCAACTAGGGCGCGCGGCCACCG 710
Db 3832 ATTGAGAATAGCTTCATGACCTCTTCTTCCAAATAGGCTCTGCCAGCTAGGG 3884

Search completed: August 17, 2003, 15:15:31
Job time : 4171.87 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 16:34:38 ; Search time 208.125 Seconds
(without alignments)
5967.841 Million cell updates/sec

Title: US-09-938-391-3

Perfect score: 555

Sequence: 1 cacaccaccagacttcca.....tgacctctcttccaagtag 555

Scoring table: OLIGO NUC

Gapop_60.0 , Gapext 60.0

Searched: 1504479 seqs, 1118970152 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0

Maximum DB-seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	529	95.3	552	12	US-10-292-418-34
2	529	95.3	552	14	US-10-131-241-50
3	392	70.6	632	14	US-10-131-241-51
4	26	4.7	534	14	US-10-042-347-6
5	26	4.7	537	14	US-10-131-241-59
6	26	4.7	546	14	US-10-042-347-4
7	26	4.7	549	12	US-10-292-418-3
8	26	4.7	549	14	US-10-131-241-53
9	26	4.7	551	13	US-10-080-797-2
10	26	4.7	552	9	US-09-873-676-30
11	26	4.7	574	14	US-10-060-036-64
12	26	4.7	900	14	US-10-131-241-45
13	26	4.7	3394	10	US-09-880-107-2178
14	26	4.7	4551	14	US-10-060-036-144
15	23	4.1	540	14	US-10-131-241-48
16	22	4.0	34	12	US-10-292-418-38

17	4.0	41	12	US-10-292-418-36	Sequence 36, Appl
18	3.4	876	14	US-10-156-761-1841	Sequence 1841, Ap
19	3.4	9025608	14	US-10-156-761-1	Sequence 1, Appli
20	3.2	430	14	US-10-198-846-11952	Sequence 11952, A
21	3.2	552	12	US-10-292-418-17	Sequence 17, Appl
22	3.2	558	9	US-09-775-174-3	Sequence 3, Appli
23	3.2	558	9	US-09-775-325-3	Sequence 3, Appli
24	3.2	565	13	US-10-036-869-37	Sequence 37, Appl
25	3.2	573	10	US-09-998-831-12	Sequence 12, Appl
26	3.2	624	13	US-10-080-797-4	Sequence 4, Appli
27	3.2	936	14	US-10-156-761-2133	Sequence 2133, Ap
28	3.2	1002	14	US-10-156-761-5807	Sequence 5807, Ap
29	3.2	17946	9	US-09-764-863-2312	Sequence 2312, Ap
30	3.2	17946	14	US-10-091-504-2312	Sequence 2312, Ap
31	3.1	399	11	US-09-918-995-12651	Sequence 12651, A
32	3.1	401	11	US-09-918-995-26500	Sequence 26500, A
33	3.1	414	13	US-10-027-632-46928	Sequence 46928, A
34	3.1	427	11	US-09-803-719-1674	Sequence 1674, Ap
35	3.1	548	9	US-09-864-761-13617	Sequence 13617, A
36	3.1	630	9	US-09-815-242-7869	Sequence 7869, Ap
37	3.1	631	13	US-10-027-632-68162	Sequence 68162, A
38	3.1	631	13	US-10-027-632-294984	Sequence 294984, A
39	3.1	638	13	US-10-027-632-102776	Sequence 102776, A
40	3.1	649	14	US-10-106-698-1578	Sequence 1578, Ap
41	3.1	837	13	US-10-027-632-153544	Sequence 153544, A
42	3.1	854	13	US-10-027-632-146821	Sequence 146821, A
43	3.1	894	14	US-10-156-761-220	Sequence 220, App
44	3.1	1083	14	US-10-156-761-1504	Sequence 1504, Ap
45	3.1	1087	9	US-09-764-870-75	Sequence 75, Appl

ALIGNMENTS

RESULT 1

US-10-292-418-34
; Sequence 34, Application US/10292418.
; Publication No. US20030139365A1
; GENERAL INFORMATION:

APPLICANT: Lo Kin-Ming
APPLICANT: Li, Yue
APPLICANT: Gillies, Stephen D
TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
FILE OF INVENTION: Immunofusins
FILE REFERENCE: LEX-006C1
CURRENT APPLICATION NUMBER: US/10/292,418
CURRENT FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: 09/383,315
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: US 60/097,883
PRIOR FILING DATE: 1998-08-25
NUMBER OF SEQ ID NOS: 54
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 34
LENGTH: 552
TYPE: DNA
ORGANISM: Canis familiaris
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(552)
OTHER INFORMATION: Endostatin
US-10-292-418-34

Query Match 95.3%; Score 529; DB 12; Length 552;

Best Local Similarity 100.0%; Pred.No. 7.1e-249; Indels 0; Gaps 0;
Matches 529; Conservative 0; Mismatches 0;

Qy 24 GGTGCTGCACCTGGTGGCCCTGAACAGCCGCGCGGCGCATCGAGGATCCGGGG 83

Db 24 GGTGCTGCACCTGGTGGCCCTGAACAGCCGCGCGGCGCATCGAGGATCCGGGG 83

Qy 84 AGCGGACTTCCAGTGTTCACAGCGCGCGCGGCGGCTGCGCGGACCTTCGGGCG 143

Db 84 AGCGACTTCAGAGTCTCCAGCAGGCGCGCGCGGGCTGGCCGGCACCTTCGCGGC 143
Qy 144 CTTCTGTGTCGCGGCTGAGGACCTCTACAGCATGTCGCGCGCGCCGACCGCACCGG 203
Db 144 CTTCTGTGTCGCGGCTGAGGACCTCTACAGCATGTCGCGCGCGCCGACCGCACCGG 203
Qy 204 GGTGCCGCTGCTAACCCTCAGGACGAGGTGCTCTTCCAGCTGGGAGGCTTATTCTC 263
Db 204 GGTGCCGCTGCTCAACCTCAGGACGAGGTGCTCTTCCAGCTGGGAGGCTTATTCTC 263
Qy 264 GGGCTCCGAGGGCCAGCTGAAGCCCGGGCCGCGATCTTCTTTTCAGCGCAGAGATGT 323
Db 264 GGGCTCCGAGGGCCAGCTGAAGCCCGGGCCGCGATCTTCTTTTCAGCGCAGAGATGT 323
Qy 324 CTTGACGACCCCGCTGCGCGCGGAGAGCGTGGCAGCGGCTCCGACCCGAGCGGGCG 383
Db 324 CTTGACGACCCCGCTGCGCGCGGAGAGCGTGGCAGCGGCTCCGACCCGAGCGGGCG 383
Qy 384 CCGCTGACCGACAGCTACTGCGAGACGTGGCGGACCGAGGCGCCCGCGCCACCGGGCA 443
Db 384 CCGCTGACCGACAGCTACTGCGAGACGTGGCGGACCGAGGCGCCCGCGCCACCGGGCA 443
Qy 444 GGGCTGCTGCTGTCGCGGCGAGGCTGCTGGAGCAGGAGGCGCGAGCTGCCGCCACGC 503
Db 444 GGGCTGCTGCTGTCGCGGCGAGGCTGCTGGAGCAGGAGGCGCGAGCTGCCGCCACGC 503
Qy 504 CTTGCTGCTGCTGTCATCGAGACAGCGTCATGACCTCTTCTCCAAG 552
Db 504 CTTGCTGCTGCTGTCATCGAGAACAGCGTCATGACCTCTTCTCCAAG 552

RESULT 2

US-10-131-241-50
; Sequence 50, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Canine sp.
US-10-131-241-50

Query Match 95.3%; Score 529; DB 14; Length 552;
Best Local Similarity 100.0%; Pred. No. 7.1e-249;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 24 GGTGCTGACCTGTTGCCCTGAACAGCCGCGAGCGGGCGGCGATGCGAGGCACTCCGGG 83
Db 24 GGTGCTGACCTGTTGCCCTGAACAGCCGCGAGCGGGCGGCGATGCGAGGCACTCCGGG 83
Qy 84 AGCGGACTTCAGAGTCTCAGCAGGCGCGCGCGGGGCTGCCCGGACCTTCCGGGC 143
Db 84 AGCGGACTTCAGAGTCTCAGCAGGCGCGCGCGGGGCTGCCCGGACCTTCCGGGC 143
Qy 144 CTTCTGTGCTGCGGCTGAGGACCTCTACAGCATGTCGCGCGCGCGGCGCGACCGG 203
Db 144 CTTCTGTGCTGCGGCTGAGGACCTCTACAGCATGTCGCGCGCGCGGCGCGACCGG 203

Qy 204 GGTGCCGCTGCTCAACCTCAGGACGAGGTGCTCTTCCAGCTGGGAGGCTTATTCTC 263
Db 204 GGTGCCGCTGCTCAACCTCAGGACGAGGTGCTCTTCCAGCTGGGAGGCTTATTCTC 263
Qy 264 GGGCTCCGAGGGCCAGCTGAAGCCCGGGCCGCGATCTTCTTTTCAGCGCAGAGATGT 323
Db 264 GGGCTCCGAGGGCCAGCTGAAGCCCGGGCCGCGATCTTCTTTTCAGCGCAGAGATGT 323
Qy 324 CTTGACGACCCCGCTGCGCGCGGAGAGCGTGGCAGCGGCTCCGACCCGAGCGGGCG 383
Db 324 CTTGACGACCCCGCTGCGCGCGGAGAGCGTGGCAGCGGCTCCGACCCGAGCGGGCG 383
Qy 384 CCGCTGACCGACAGCTACTGCGAGACGTGGCGGACCGAGGCGCCCGCGCCACCGGGCA 443
Db 384 CCGCTGACCGACAGCTACTGCGAGACGTGGCGGACCGAGGCGCCCGCGCCACCGGGCA 443
Qy 444 GGGCTGCTGCTGTCGCGGCGAGGCTGCTGGAGCAGGAGGCGCGAGCTGCCGCCACGC 503
Db 444 GGGCTGCTGCTGTCGCGGCGAGGCTGCTGGAGCAGGAGGCGCGAGCTGCCGCCACGC 503
Qy 504 CTTGCTGCTGCTGTCATCGAGAACAGCGTCATGACCTCTTCTCCAAG 552
Db 504 CTTGCTGCTGCTGTCATCGAGAACAGCGTCATGACCTCTTCTCCAAG 552

RESULT 3

US-10-131-241-51
; Sequence 51, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
; LENGTH: 632
; TYPE: DNA
; ORGANISM: Murinae sp.
US-10-131-241-51

Query Match 70.6%; Score 392; DB 14; Length 632;
Best Local Similarity 100.0%; Pred. No. 6.4e-182;
Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 161 TCGAGGACCTCTACAGCATGTCGCGCGCGCGAGCCGCGGTCGCCGCTCGTCAACC 220
Db 241 TCGAGGACCTCTACAGCATGTCGCGCGCGCGAGCCGCGGTCGCCGCTCGTCAACC 300
Qy 221 TCAGGACGAGGTGCTTCTCCAGCTGGGAGGCTTATTCTCGGGCTCCGAGGCGCAG 280
Db 301 TCAGGACGAGGTGCTTCTCCAGCTGGGAGGCTTATTCTCGGGCTCCGAGGCGCAG 360
Qy 281 TGAAGCCCGGGCCCGCATCTTCTTTTCAGCGGAGAGATGCTCTGAGCACCCTGCT 340
Db 361 TGAAGCCCGGGCCCGCATCTTCTTTTCAGCGGAGAGATGCTCTGAGCACCCTGCT 420
Qy 341 GGCCCGGAGAGCGTGTGGACGCGCTCCGACCCGAGCGGGCGCGCTGACCGCAGCT 400
Db 421 GGCCCGGAGAGCGTGTGGACGCGCTCCGACCCGAGCGGGCGCGCTGACCGCAGCT 480
Qy 401 ACTGCGAGACGTGGCGGACGAGGCGCCCGGCGGCCACCGGCGAGGCGTCTGCTGCTCG 460

; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: US 60/097,883
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(549)
; OTHER INFORMATION: endostatin
US-10-292-418-3

Query Match 4.7%; Score 26; DB 12; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.0052;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGCTTCCAGCAGCGCG 113
|||||
Db 88 GACTTCCAGTGCTTCCAGCAGCGCG 113

RESULT 8

US-10-131-241-53
; Sequence 53, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-241-53

Query Match 4.7%; Score 26; DB 14; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.0052;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGCTTCCAGCAGCGCG 113
|||||
Db 88 GACTTCCAGTGCTTCCAGCAGCGCG 113

RESULT 9

US-10-080-797-2
; Sequence 2, Application US/10080797
; Publication No. US20020183253A1
; GENERAL INFORMATION:
; APPLICANT: Campochiaro, Peter A.
; APPLICANT: Dixon, Katherine H.
; APPLICANT: Brazzell, Romulus K.
; TITLE OF INVENTION: METHOD FOR TREATING OCULAR
; TITLE OF INVENTION: NEOVASCULARIZATION
; FILE REFERENCE: 4-31881A
; CURRENT APPLICATION NUMBER: US/10/080,797
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 21

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Human
US-10-080-797-2

Query Match 4.7%; Score 26; DB 13; Length 551;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGCTTCCAGCAGCGCG 113
|||||
Db 87 GACTTCCAGTGCTTCCAGCAGCGCG 112

RESULT 10

US-09-873-676-30
; Sequence 30, Application US/09873676
; Patent No. US2002007289A1
; GENERAL INFORMATION:
; APPLICANT: MacDonald, Nicholas J.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use
; FILE REFERENCE: 05213-0378 (43170-259333)
; CURRENT APPLICATION NUMBER: US/09/873,676
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-676-30

Query Match 4.7%; Score 26; DB 9; Length 552;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGCTTCCAGCAGCGCG 113
|||||
Db 88 GACTTCCAGTGCTTCCAGCAGCGCG 113

RESULT 11

US-10-060-036-64
; Sequence 64, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-036-64

Query Match 4.7%; Score 26; DB 14; Length 574;

Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGTCTCCAGCAGCGCG 113
Db 422 GACTTCCAGTGTCTCCAGCAGCGCG 447
|||||

RESULT 12
US-10-131-241-45
; Sequence 45, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-241-45

Query Match 4.7%; Score 26; DB 14; Length 900;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGTCTCCAGCAGCGCG 113
Db 148 GACTTCCAGTGTCTCCAGCAGCGCG 173
|||||

RESULT 13
US-09-880-107-2178
; Sequence 2178, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2178
; LENGTH: 3394
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L22548
US-09-880-107-2178

Query Match 4.7%; Score 26; DB 10; Length 3394;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGTCTCCAGCAGCGCG 113
Db 1591 GACTTCCAGTGTCTCCAGCAGCGCG 1616
|||||

RESULT 14
US-10-060-036-144
; Sequence 144, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yudi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT FILING DATE: 2002-01-30
; CURRENT APPLICATION NUMBER: US/10/060,036
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144
; LENGTH: 4551
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-036-144

Query Match 4.7%; Score 26; DB 14; Length 4551;
Best Local Similarity 100.0%; Pred. No. 0.0035;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGTCTCCAGCAGCGCG 113
Db 4087 GACTTCCAGTGTCTCCAGCAGCGCG 4112
|||||

RESULT 15
US-10-131-241-48
; Sequence 48, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 540
; TYPE: DNA
; ORGANISM: Rhesus monkey
US-10-131-241-48

Query Match 4.1%; Score 23; DB 14; Length 540;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGTCTCCAGCAGCG 110
Db 88 GACTTCCAGTGTCTCCAGCAGCG 110
|||||

Search completed: August 17, 2003, 19:58:14
Job time : 214.125 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 16:23:08 ; Search time 2416.32 Seconds
(without alignments)
8338.472 Million cell updates/sec

Title: US-09-938-391-1
Perfect score: 829
Sequence: 1 cctgsgcggcagatgacat.....aataaaggagcgaag 829

Scoring table: OLIGO NUC
Gapop_60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vxl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	47	5.7	279	9 AW464343	AW464343 BP230015B
2	47	5.7	387	10 BF601253	BF601253 266182 MA
3	47	5.7	551	10 BF074459	BF074459 221883 MA
4	41	4.9	657	14 CB444165	CB444165 695295 MA

35	4.2	110	10	BF322807	BF322807 maa33hl2.
35	4.2	120	10	BGI51469	BGI51469 nai38a11.
35	4.2	138	12	BM197778	BM197778 C0324C07-
35	4.2	142	9	AI540493	AI540493 tq25a01.x
35	4.2	155	13	BU610789	BU610789 UT-M-FCO-
35	4.2	158	9	AI991894	AI991894 w442a08.x
35	4.2	162	12	BI327804	BI327804 602979552
35	4.2	171	9	AI561355	AI561355 tq48e04.x
35	4.2	206	9	AV112006	AV112006 AV112006
35	4.2	215	12	BM148522	BM148522 TCAAP2D13
35	4.2	239	10	BB066072	BB066072 BB066072
35	4.2	251	9	AA292773	AA292773 z156b11.s
35	4.2	255	9	AI964083	AI964083 wt88g11.x
35	4.2	267	13	BU070539	BU070539 UT-M-FRO-
35	4.2	268	10	BF057259	BF057259 7k18b04.x
35	4.2	268	10	BF223272	BF223272 7q31e04.x
35	4.2	282	9	AI932582	AI932582 w04e06.x
35	4.2	288	9	AA295509	AA295509 EST100752
35	4.2	295	13	BY671598	BY671598 BY671598
35	4.2	296	13	BY367100	BY367100 BY367100
35	4.2	298	12	BI144536	BI144536 602909058
35	4.2	302	9	AI223972	AI223972 qx07h03.x
35	4.2	305	9	AW513899	AW513899 x90b01.x
35	4.2	305	10	BF061815	BF061815 7k67d01.x
35	4.2	307	9	AW136127	AW136127 UI-H-B11-
35	4.2	321	10	BB039192	BB039192 BB039192
35	4.2	335	9	AL047369	AL047369 DKF2p586K
35	4.2	340	9	AI872536	AI872536 w09e03.x
35	4.2	349	13	BQ233745	BQ233745 AGENCOURT
35	4.2	357	10	BF436805	BF436805 7p64c10.x
35	4.2	358	14	CD540078	CD540078 B0219D01-
35	4.2	365	9	AI217549	AI217549 qd43b10.x
35	4.2	366	13	BY674034	BY674034 BY674034
35	4.2	370	10	BB731921	BB731921 BB731921
35	4.2	372	9	AI633945	AI633945 tt23d06.x
35	4.2	373	10	BB731964	BB731964 BB731964
35	4.2	374	10	BB731918	BB731918 BB731918
35	4.2	374	13	BY660494	BY660494 BY660494
35	4.2	375	9	AI241051	AI241051 qj95g10.x
35	4.2	378	10	BB731968	BB731968 BB731968
35	4.2	378	13	BY670329	BY670329 BY670329

ALIGNMENTS

RESULT 1
AW464343 279 bp mRNA linear EST 24-FEB-2000
LOCUS BP230015B10C11 Soares normalized bovine placenta Bos taurus CDNA
DEFINITION clone BP230015B10C11 5', mRNA sequence.

ACCESSION AW464343.1 GI:7034511

VERSION EST.

KEYWORDS Bos taurus (cow)

SOURCE Bos taurus

ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 279)

AUTHORS Lewin,H.A., Soares,M.B., Rebeiz,M., Pardini,J., Liu,L. and Larson

J.H.

TITLE Bovine ESTs

JOURNAL Unpublished

COMMENT Contact: Lewin, H. A.

W. M. Keck Center for Comparative and Functional Genomics

University of Illinois at Urbana-Champaign

340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL

61801, USA

Tel: 217 333 5998

Fax: 217 244 5617

Email: h-lewin@uiuc.edu

Funding for cattle EST sequencing was provided by the USDA National

Research Initiative, Animal Genome Resource Grant AG 99-3205-8534 to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED from Washington University Genome Center. Vector Trimmi g: Cross match from Washington University Genome Center PHRAP suite. Sequences submitted are vector free and at least 200 bp in length.

PCR Primers

FORWARD: TAATACGACTCACTATAGG

BACKWARD: ATTAACCTCACTAAG

Insert Length: 279 Std Error: 0.00

Plate: BP230015B10 row: C column: 11

Seq primer: AGCGATAACAATTACACAGGA

High quality sequence stop: 279.

Location/Qualifiers

1. .279

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/clone="BP230015B10C11"

/sex="female"

/lab_host="DH10B"

/clone_lib="Soares normalized bovine placenta"

/note="Organ: placenta; Vector: pT73Pac; Site 1: EcoRI;

Site 2: NotI; The cDNA library was contributed by the

Soares laboratory and it was constructed and normalized

as described by Bonaldo, M.F., Lennon, G. and Soares,

M.B. (1996), Genome Research 6(9): 791-806."

35 a 106 c 86 g 52 t

BASE COUNT

ORIGIN

Query Match 5.7%; Score 47; DB 9; Length 279;

Best Local Similarity 100.0%; Pred. No. 7.9e-12;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 232 CAGTGTTCAGCAGCGCGCGCGGCTGGCGGACCTTCG 278

|||||

Db 24 CAGTGTTCAGCAGCGCGCGCGGCTGGCGGACCTTCG 70

RESULT 2

BF601253

LOCUS

BF601253 387 bp mRNA linear EST 25-APR-2001

DEFINITION 266182 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION BF601253

VERSION BF601253.1 GI:11698475

KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;

Bovidae; Bovinae; Bos.

1 (bases 1 to 387)

Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,

Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett

, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,

Perteau, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and

Keele, J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA

libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

21180013

11282978

COMMENT

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 18

and -mismatch 12 options.

PCR Primers

FORWARD: AGGAAACAGCTATGACCAT

BACKWARD: GTTTCCAGTCACGACG

Plate: 40 row: H column: 7

Seq primer: ATTAGTGCACACTATAG.

FEATURES

source

1. .387

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/tissue_type="pooled"

/lab_host="DH10B"

/clone_lib="MARC 3BOV"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;

Library made from pooled tissue from marrow, alveolar

macrophage, ovary, fetal semitendinosus muscle, and fetal

longissimus muscle."

52 a 140 c 131 g 64 t

BASE COUNT

ORIGIN

Query Match 5.7%; Score 47; DB 10; Length 387;

Best Local Similarity 100.0%; Pred. No. 8.1e-12;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 232 CAGTGTTCAGCAGCGCGCGCGGCTGGCGGACCTTCG 278

|||||

Db 7 CAGTGTTCAGCAGCGCGCGCGGCTGGCGGACCTTCG 53

RESULT 3

BF074459

LOCUS

BF074459 551 bp mRNA linear EST 25-APR-2001

DEFINITION 221883 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION BF074459

VERSION BF074459.1 GI:10867970

KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;

Bovidae; Bovinae; Bos.

1 (bases 1 to 551)

Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,

Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett

, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,

Perteau, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and

Keele, J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA

libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

21180013

11282978

COMMENT

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 18

and -mismatch 12 options.

PCR Primers

FORWARD: AGGAAACAGCTATGACCAT

BACKWARD: GTTTCCAGTCACGACG

Plate: 81 row: F column: 9

Seq primer: ATTAGTGCACACTATAG.

FEATURES

source

1. .551

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/tissue_type="pooled"

/lab_host="DH10B"

/clone_lib="MARC 2BOV"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;

Library made from pooled tissue from testis, thymus,

found through the I.M.A.G.E. Consortium/LLNL, send email to:

info@image.llnl.gov
 Putative full length read
 The vector to vector length is 121
 Seq primer: -40RP from Gibco.
 Location/Qualifiers
 1. 120
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4262301"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP HN20"
 /notes="Organ: normal_head/neck tissue; Vector: pAMP1; mRNA
 made from head/neck tissue, cDNA made by oligo-dT
 priming. Directionally cloned into UDG sites.
 Size-selected on agarose gel, average insert size 300 bp.
 Primary library. cDNA Library Preparation: David B.
 Krizman, Ph.D."

BASE COUNT 44 a 24 c 22 g 30 t
 ORIGIN
 Query Match 4.2%; Score 35; DB 10; Length 120;
 Best Local Similarity 100.0%; Pred. No. 5.2e-06;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCTCTCAAGAAATAAAGAGCCCAAGAG 829

Db 60 ATGTAATCTCTCAAGAAATAAAGAGCCCAAGAG 94

RESULT 7
 BM197778/c
 LOCUS BM197778 138 bp mRNA linear EST 08-JUN-2003
 DEFINITION Mus musculus cDNA clone NIA: C0324C07 IMAGE:30007134 3', mRNA

ACCESSION BM197778
 VERSION BM197778.2 GI:31522625
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 138)
 Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
 Construction of long-transcript enriched cDNA libraries from
 submicrogram amounts of total RNAs by a universal PCR amplification
 method
 Genome Res. 11 (9), 1553-1558 (2001)
 21429098
 11544199

On Dec 14, 2001 this sequence version replaced gi:17750067.

Other ESTs: C0324C07-SN
 Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 Plate: C0324 row: C column: 07
 Seq primer: -21M13 Forward
 High quality sequence stop: 138
 POLYA=Yes.

FEATURES
 source
 1. 138
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129/Sv x 129/Sv-CP"
 /db_xref="niaEST:C0324C07-3"
 /db_xref="taxon:10090"
 /clone="NIA:C0324C07 IMAGE:30007134"
 /tissue_type="Undifferentiated ES Cell"
 /cell_line="R1 ES cells"

/lab_host="DH10B"
 /clone_lib="NIA Mouse Undifferentiated ES Cell cDNA
 Library (Long)"
 /note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
 NotI; Mouse cDNA project by the Laboratory of Genetics,
 National Institute on Aging (NIA), Intramural Research
 Program, NIH (<http://lgsun.grc.nia.nih.gov/cDNA>). This is a
 long-transcript enriched cDNA library (Ref. Genome Res.
 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
 obtained from Dr. Kenneth R. Boheler (National Institute
 on Aging, USA). ES cells were cultured without feeder
 cells in the presence of LIF and BRL-conditioned media.
 Double-stranded cDNAs were synthesized with an Oligo(dT)
 primer (Invitrogen).
 5'-pGACTAGTCTAGATCGGCGCGCCCTTTT-3') from
 14.2 ug of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to Loner-linker LL-Sal4, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Taq polymerase (Takara) with a primer Sal4-S. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with SalI and NotI enzymes
 and cloned into SalI/NotI site of pSPORT1 plasmid vector.
 The DH10B E. coli host was transformed with the ligation
 mixture by the standard chemical method. The average
 insert size is about 2.4 kb. The library was constructed
 by Yulan Piao (NIA)." 29 a 35 c 31 g 43 t
 BASE COUNT 29 a 35 c 31 g 43 t
 ORIGIN

Query Match 4.2%; Score 35; DB 12; Length 138;
 Best Local Similarity 100.0%; Pred. No. 5.3e-06;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 795 ATGTAATCTCTCAAGAAATAAAGAGCCCAAGAG 829
 Db 36 ATGTAATCTCTCAAGAAATAAAGAGCCCAAGAG 2

RESULT 8
 AI540493/c
 LOCUS AI540493 142 bp mRNA linear EST 13-MAY-1999
 DEFINITION tG25a01.x1 NCI_CGAP_Ut1 Homo sapiens cDNA clone IMAGE:2209800 3',
 mRNA sequence.
 ACCESSION AI540493
 VERSION AI540493.1 GI:4457866
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 142)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 716 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 113
 POLYA=No.

FEATURES
 source
 1. 142
 Location/Qualifiers

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2209800"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Utl1"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Life Technologies catalog #:
11538-014"

BASE COUNT      24 a  40 c  42 g  35 t      1 others
ORIGIN
Query Match      4.2%; Score 35; DB 9; Length 142;
Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795' ATGTAATCCTCAAGAAATAAAAGGAGCCCAAGAG 829
|||||
Db 39 ATGTAATCCTCAAGAAATAAAAGGAGCCCAAGAG 5

RESULT 9
BU610789      155 bp mRNA linear EST 20-FEB-2003
LOCUS
DEFINITION
UI-M-FCO-car-h-11-0-UI-rl NIH BMAP FCO Mus musculus cDNA clone
BU610789
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)

REFERENCE
1 (bases 1 to 155)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
The following repetitive elements were found in this cDNA
sequence: 103-145, >POLY_A#simple_repeat
Seq primer: pYX-5,
Location/Qualifiers
1. .155
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="UI-M-FCO-car-h-11-0-UI"
/tissue_type="whole brain"
/dev_stage="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FCO"
/notes="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
, is TGAGAGAGCC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT      59 a  33 c  32 g  31 t
ORIGIN
Query Match      4.2%; Score 35; DB 13; Length 155;
Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCCTCAAGAAATAAAAGGAGCCCAAGAG 829
|||||
Db 92 ATGTAATCCTCAAGAAATAAAAGGAGCCCAAGAG 126

RESULT 10
AI991894      158 bp mRNA linear EST 08-MAR-2000
LOCUS
DEFINITION
ws42a08.x1 NCI CGAP Brn25 Homo sapiens cDNA clone IMAGE:2499830 3'
similar to contains TARI.tl MER22 repetitive element 1, mRNA
sequence.
AI991894
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)

REFERENCE
1 (bases 1 to 158)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BTGAP), Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1209 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 144.
Location/Qualifiers
1. .158
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2499830"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Brn25"
/notes="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCCGCGATAGGTTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT      26 a  50 c  45 g  37 t
ORIGIN
Query Match      4.2%; Score 35; DB 9; Length 158;

```

```

Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCTCTCAGAAATAAAGGAAGCAAGAG 829
|||||
Db 38 ATGTAATCTCTCAGAAATAAAGGAAGCAAGAG 4

RESULT 11
BI327804 162 bp mRNA linear EST 30-JUL-2001
LOCUS 602979552F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:5132580 5',
DEFINITION mRNA sequence.
ACCESSION BI327804
VERSION BI327804.1 GI:15012461
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 162)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Arrayed by: Life Technologies, Inc.
CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Distribution: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11325 row: 9 column: 13
High quality sequence stop: 160.
FEATURES
source
1..162
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="FVB/N"
/clone="IMAGE:5132580"
/lab_host="NCI CGAP Li9"
/clone_lib="NCI CGAP Li9"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 61 a 33 c 41 g 27 t
ORIGIN
Query Match 4.2%; Score 35; DB 12; Length 162;
Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCTCTCAGAAATAAAGGAAGCAAGAG 829
|||||
Db 111 ATGTAATCTCTCAGAAATAAAGGAAGCAAGAG 145

RESULT 12
AI561355/c 171 bp mRNA linear EST 24-MAR-1999
LOCUS tq48e04.x1 NCI_CGAP_Utl1 Homo sapiens cDNA clone IMAGE:221262 3',
DEFINITION mRNA sequence.
ACCESSION AI561355
VERSION AI561355.1 GI:4511696
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 171)

AUTHORS
TITLE
JOURNAL
COMMENT
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 148.
FEATURES
source
1..171
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:221262"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP Utl1"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Life Technologies catalog #:
11538-014"
BASE COUNT 35 a 50 c 45 g 41 t
ORIGIN
Query Match 4.2%; Score 35; DB 9; Length 171;
Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCTCTCAGAAATAAAGGAAGCAAGAG 829
|||||
Db 54 ATGTAATCTCTCAGAAATAAAGGAAGCAAGAG 20

RESULT 13
AV112006 206 bp mRNA linear EST 29-JUN-1999
LOCUS AV112006 Mus musculus C57BL/6J 10-day embryo Mus musculus CDNA
DEFINITION Clone 2610008A07, mRNA sequence.
ACCESSION AV112006
VERSION AV112006.1 GI:5266086
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 206)
Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,
Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara
A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J.,
Kikuchi, N., Kojima, Y., Matsuyama, T., Naitsumu, H., Oda, H., Owa, C.,
Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara
Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y., Tomimaga, N.,
Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M.,
Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
RIKEN Mouse ESTs
Unpublished
Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@tc.riken.go.jp
Thermotabilization and thermoactivation of thermolabile enzymes by

```


trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES

source
1. .206
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="261008A07"
/sex="mixed"
/dev_stage="10-day embryo"
/clone_lib="Mus musculus C57BL/6J 10-day embryo"

BASE COUNT 54 a 39 c 50 g 63 t

ORIGIN

Query Match 4.2%; Score 35; DB 9; Length 206;
Best Local Similarity 100.0%; Pred. No. 5.4e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCTCTCAAGAAATAAAGGAGCCAAAGAG 829

Db 161 ATGTAATCTCTCAAGAAATAAAGGAGCCAAAGAG 195

RESULT 14

BM148522/c

LOCUS

DEFINITION BM148522 215 bp mRNA linear EST 30-NOV-2001
TCAAP2D13934 Pediatric acute myelogenous leukemia cell (FAB M1)
Baylor-HGSC project=TCAA Homo sapiens cDNA clone TCAAP1393, mRNA sequence.

ACCESSION

BM148522 GI:17169320

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 215)

Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, F.R. Jr.,

Gunaratne, P.H., Muzny, D., Bouck, J.J., Gibbs, R.A. and Margolin, J.F.

Pediatric Leukemia cDNA Sequencing Project (2001)

Unpublished

Contact: Dr. Judith F. Margolin

Texas Children's Cancer Center and Human Genome Sequencing Center

at Baylor College of Medicine

1102 Bates, MC3-3320 Houston, TX 77030, USA

Tel: 832-824-4536

Fax: 832-825-4038

Email: clones@ccc.org

Seq primer: M13 primer.

Location/Qualifiers

1. .215

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="TCAAP1393"

/sex="male"

/tissue_type="leukopheresis"

/cell_type="myeloid cell"

/dev_stage="pediatric 6 years"

/lab_host="DH10B"

/clone_lib="Pediatric acute myelogenous leukemia cell (FAB

M1) Baylor-HGSC project=TCAA"

/notes="Vector: lambda pSB; Site 1: BamHI; Site 2: EcoRI;

First strand cDNA was primed with an anchored

XhoI-oligo(dT) primer [5'CGAGACTCGAGCGCGCAGGAG(T)VN

3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand

was primed with a BamHI-dC primer

[5'AGAGAGCTCGATCCGCGCGCAATAATAAT(C) 3'].

Double-stranded cDNA was then digested with BamHI and XhoI and directionally cloned into the BamHI and SalI sites of lambda pSB vector. Library went through one round of normalization. Library was constructed by Wei Yu at RIKEN of Japan (Carninci P., Westover A., Nishiyama Y., Ohsumi T., Itoh M., Nagaoka S., Sasaki, Okazaki Y., Muramatsu M., Schneider C., Hayashizaki Y., High efficiency selection of full-length cDNA by improved biotinylated cap trapper., DNA Res 4: 1, 61-6, Feb 28, 1997)"

BASE COUNT 38 a 54 c 47 g 76 t

ORIGIN

Query Match 4.2%; Score 35; DB 12; Length 215;
Best Local Similarity 100.0%; Pred. No. 5.4e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCTCTCAAGAAATAAAGGAGCCAAAGAG 829

Db 84 ATGTAATCTCTCAAGAAATAAAGGAGCCAAAGAG 50

RESULT 15

BB066072

LOCUS

DEFINITION

BB066072 239 bp mRNA linear EST 25-JUN-2000

BB066072 RIKEN full-length enriched, 15 days embryo male testis Mus

musculus cDNA clone 8030447L20 3' similar to D17546 Mouse mRNA for

collagen, mRNA sequence.

ACCESSION

BB066072

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 239)

Komno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci

, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,

Izawa, M., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,

Kawano, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,

Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.

, Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.

, Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata

, Y., Shigenoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y.

, Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya

, T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaka, I.

, Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino

, M., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Konno, H., et al.)

Unpublished

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp

URL: <http://genome.gsc.riken.go.jp/>

Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki

, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Thermotabilization and thermoactivation of thermostable enzymes by

trehalose and its application for the synthesis of full length

cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)

Itoh, M., Katsunari, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,

Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki

, Y. and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation

system. Genome Res. 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning. Methods Enzymol. 303,

19-44 (1999)

Please visit our web site (<http://genome.rtc.riken.go.jp>) for

further details.

FEATURES
source

```

Location/Qualifiers
1. .239
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/cdname="8030447L20"
/sex="male"
/tissue_type="testis"
/dev_stage="15 days embryo"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 15 days embryo
male testis"
/notes="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5']
GAGAGAGAGAGATCCACAGACTCTTTTCTTTTTTTTTTN 3']. cDNA was
prepared by using trihalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of subtraction to
Rot = 185.0. Second strand cDNA was prepared with the
primer adapter of sequence [5']
GAGAGAGAGATCTCGAGTTATTAATAATCCCCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI"
74 C 47 G 55 T

```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 15:15:38 ; Search time 4161.77 Seconds
(without alignments)
8148.953 Million cell updates/sec

Title: US-09-938-391-1
Perfect score: 829
Sequence: 1 cccctggcggcagatgacat.....aataaaaggagcccaagag, 829

Scoring table: OLIGO.NUC
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.ste.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.ste.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_man.*
- 37: em.htg_vrt.*
- 38: em.sv.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	829	100.0	829	6	AX399629	AX399629 Sequence
2	555	66.9	555	6	AX399631	AX399631 Sequence
C 3	35	4.2	251	6	AX408066	AX408066 Sequence
C 4	35	4.2	338	6	AX524393	AX524393 Sequence
C 5	35	4.2	338	6	AX553131	AX553131 Sequence
6	35	4.2	1615	10	MMU03714	U03714 Mus musculus
7	35	4.2	2364	9	AK098216	AK098216 Homo sapi
8	35	4.2	2637	9	BC033715	BC033715 Homo sapi
9	35	4.2	3129	10	MMCOL18A11	U03715 Mus musculus
10	35	4.2	3394	6	AX409531	AX409531 Sequence
11	35	4.2	3394	6	I51045	I51045 Sequence 4
12	35	4.2	3394	9	HUMCOL18AX	L22548 Human colla
13	35	4.2	4031	6	I51044	I51044 Sequence 1
14	35	4.2	4031	10	MUSCOLPRO	L22545 Mouse alpha
15	35	4.2	4437	10	MUSCOLLAG	D17546 Mus musculus
16	35	4.2	4802	10	BC043697	BC043697 Mus muscu
17	35	4.2	5408	9	AF018082	AF018082 Homo sapi
18	35	4.2	5929	9	AF018081	AF018081 Homo sapi
19	35	4.2	131563	9	HSE12310	BX322561 Homo sapi
20	35	4.2	186105	10	AC055777	AC055777 Mus muscu
21	35	4.2	188294	2	AC007641	AC007641 Mus muscu
C 22	35	4.2	223201	9	HS53110	AL133493 Homo sapi
23	35	4.2	340000	9	HS21C102	AL163302 Homo sapi
24	28	3.4	851	10	AF189709	AF189709 Rattus no
25	27	3.3	900	6	AX370851	AX370851 Sequence
26	26	3.1	534	6	AR193166	AR193166 Sequence
27	26	3.1	537	6	AX100092	AX100092 Sequence
28	26	3.1	546	6	AR193165	AR193165 Sequence
29	26	3.1	549	6	AX100086	AX100086 Sequence
30	26	3.1	552	6	AX395662	AX395662 Sequence
31	26	3.1	552	6	AX473835	AX473835 Sequence
32	26	3.1	552	6	BD081407	BD081407 Fused pro
33	26	3.1	555	9	AF184060	AF184060 Homo sapi
34	26	3.1	564	9	AF416592	AF416592 Homo sapi
35	26	3.1	650	6	AX370853	AX370853 Sequence
36	26	3.1	786	9	AF282883	AF282883 Homo sapi
37	26	3.1	238671	2	AC125296	AC125296 Rattus no
38	25	3.0	183076	10	AL807386	AL807386 Mouse DNA
39	24	2.9	67	8	AF163969	AF163969 Saccharom
C 40	24	2.9	36775	3	AC078947	AC078947 Leishmani
41	24	2.9	59538	2	AC012951	AC012951 Drosophil
42	24	2.9	104416	2	AC020324	AC020324 Drosophil
C 43	24	2.9	139597	3	LMFLUNK01	AL356456 Leishmani
44	24	2.9	158519	9	AL135926	AL135926 Human DNA
C 45	24	2.9	164719	3	AC009374	AC009374 Drosophil

ALIGNMENTS

RESULT 1	AX399629	Sequence	1	from Patent	EP1191036..	829 bp	DNA	linear	PAT 06-JUN-2002
LOCUS	AX399629	Sequence	1	from Patent	EP1191036..	829 bp	DNA	linear	PAT 06-JUN-2002
DEFINITION	AX399629	Sequence	1	from Patent	EP1191036..	829 bp	DNA	linear	PAT 06-JUN-2002
ACCESSION	AX399629	Sequence	1	from Patent	EP1191036..	829 bp	DNA	linear	PAT 06-JUN-2002
VERSION	AX399629.1	GI:21335409							
KEYWORDS	Canis familiaris (dog)								
SOURCE	Canis familiaris								
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.								
REFERENCE	1	Sheppard, M.G. and Tong, X.							
AUTHORS	Methods and compositions for diagnosing and treating disorders								
TITLE	involving angiogenesis								


```
VERSION AX408066.1 GI:21440771
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL
PATENT: WO 0229103-A 713 11-APR-2002;
GENE EXPRESSION PROFILES IN LIVER CANCER
GENE LOGIC INC (US)
FEATURES
source
1. .251
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/notes="EMBL/GenBank Accession No. AA292773"
BASE COUNT 44 a 72 c 75 g 60 t
ORIGIN
Query Match 4.2%; Score 35; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCCTCAAGAAATAAAGGAAGCCCAAGAG 829
Db 41 ATGTAATCCTCAAGAAATAAAGGAAGCCCAAGAG 7

RESULT 4
AX524393/c
LOCUS AX524393 338 bp DNA linear PAT 21-NOV-2002
DEFINITION Sequence 423 from Patent EP1236798.
ACCESSION AX524393
VERSION AX524393.1 GI:25169489
KEYWORDS Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL
PATENT: EP 1236798-A 423 04-SEP-2002;
LION BIOSCIENCE AG (DE)
FEATURES
source
1. .338
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
BASE COUNT 71 a 92 c 83 g 92 t
ORIGIN
Query Match 4.2%; Score 35; DB 6; Length 338;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCCTCAAGAAATAAAGGAAGCCCAAGAG 829
Db 39 ATGTAATCCTCAAGAAATAAAGGAAGCCCAAGAG 5

RESULT 5
AX553131/c
LOCUS AX553131 338 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 423 from Patent WO02074953.
ACCESSION AX553131
VERSION AX553131.1 GI:25897131
KEYWORDS Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 Hofer, M., Hofmann, M., Kaiser, C., Kranz, H., Loebbert, R. and
Schlueter, T.
TITLE Gene library and a method for producing the same
JOURNAL Patent: WO 02074953-A 423 26-SEP-2002;
LION BIOSCIENCE AG (DE)
FEATURES
source
1. .338
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
BASE COUNT 71 a 92 c 83 g 92 t
ORIGIN
Query Match 4.2%; Score 35; DB 6; Length 338;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCCTCAAGAAATAAAGGAAGCCCAAGAG 829
Db 39 ATGTAATCCTCAAGAAATAAAGGAAGCCCAAGAG 5

RESULT 6
MMU03714
LOCUS MMU03714 1615 bp mRNA linear ROD 13-AUG-1994
DEFINITION partial cds.
ACCESSION U03714
VERSION U03714.1 GI:487733
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 Rehn, M.V. and Pihlajaniemi, T.
TITLE A previously uncharacterized collagen chain, alpha 1(XVIII), with
frequent interruptions in the collagenous sequence, a distinct
tissue distribution, and homology with type XV collagen
JOURNAL Proc. Natl. Acad. Sci. U.S.A. (1994) In press
AUTHORS Rehn, M., Hintikka, E. and Pihlajaniemi, T.
TITLE Primary structure of the alpha 1 chain of mouse type XVIII
collagen, partial structure of the corresponding gene, and
comparison of the alpha 1(XVIII) chain with its homologue, the
alpha 1(XV) collagen chain
JOURNAL J. Biol. Chem. 269 (19), 13929-13935 (1994)
MEDLINE 94245707
PUBMED 8188673
REFERENCE
AUTHORS Rehn, M.V.
TITLE Direct Submission
JOURNAL Submitted (24-NOV-1993) Marko Rehn, Department of Medical
Biochemistry, University of Oulu, Kajaanintie 52 A Fin-90220, Oulu,
Finland
FEATURES
source
1. .1615
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/sex="male"
/tissue_type="liver"
/clone_lib="Mouse adult liver cDNA library by M. Rehn"
/dev_stage="adult, 9-10 weeks"
<1. _1449
/standard_name="COL18A1"
/citation=[2]
/citation=[1]
/codon_start=1
/product="alpha 1(XVIII) collagen"
CDS
```

```

/protein_id="AAA20657.1"
/db_xref="GI:487734"
/translation="DNAFVESCRLGLPGQGVQSPGKDGKGEVPPGPPQCPQPID
LFILREMKDGRDAGQGERGEPFAGPGGFFSSVFPGPFGYGIPIGPKGES
IPGPPGPPGQGGPIGYEGRQGGPPGPPGPPGPPGPHRQTVSVGPPGPPGPP
PGMGASAGQVRWIWYQTMIDKIRVEGWLIFVAERELYYVRNGRFRKVLLEART
ALPRGTGNEVAQLQPELVOLHGEPSYTRREYSYSTARPWRADILANRPLPRQFPY
GVPHHSSYVHLPPARTLSLAHTHODFOPLVHLVALNTPLSGMGIIGADQFCFQQ
ARAVGISCFERAFLSRLDLYSIVRADRGSPYIVNLKDEVLSPSWDSIFSGSQQL
QPGARISFDGRVLRHPAWPKQSVWHGSDPSGRRLMESYCYTWTTRTGTATGQASSL
LSGRLLQEKAASCHNSYIVLCIENSFMTSPSK"
/citation=[2]
/citation=1584..1589
polya_signal
BASE COUNT 338 a 493 c 479 g 305 t
ORIGIN
Query Match 4.2%; Score 35; DB 10; Length 1615;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 795 ATGTAATCTCTCAAGAAATAAAAGAACCCAAAGAG 829
Db 1569 ATGTAATCTCTCAAGAAATAAAAGAACCCAAAGAG 1603
RESULT 7
AK098216 2364 bp mRNA linear PRI 15-JUL-2002
LOCUS Homo sapiens cDNA FLJ40897 fis, clone UTERU2002964, highly similar
DEFINITION to Homo sapiens type XVIII collagen (COL18A1) mRNA.
ACCESSION AK098216
VERSION AK098216.1 GI:21758185
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K.,
Arita, M., Musashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B.,
Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and
Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2364)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
1..2364
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UTERU2002964"
/tissue_type="uterus"
/clone_lib="UTERU2"
/notes="cloning vector: pME189FL3"
FEATURES
source
1..2364
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UTERU2002964"
/tissue_type="uterus"
/clone_lib="UTERU2"
/notes="cloning vector: pME189FL3"
CDS
FEATURES
Location/Qualifiers
1..2637
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4425380"
/tissue_type="Kidney, hypernephroma"
/clone_lib="NIH MGC_89"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
<1..2453
/codon_start=3
/product="Similar to collagen, type XVIII, alpha 1"
/protein_id="AAH33715.1"
/db_xref="GI:21708149"
/translation="TRPPPGREGPPGFLGPGPPGPPGREGPPGRTGKGLGEAGAP
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 68 Row: e Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
Location/Qualifiers
1..2637
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4425380"
/tissue_type="Kidney, hypernephroma"
/clone_lib="NIH MGC_89"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
<1..2453
/codon_start=3
/product="Similar to collagen, type XVIII, alpha 1"
/protein_id="AAH33715.1"
/db_xref="GI:21708149"
/translation="TRPPPGREGPPGFLGPGPPGPPGREGPPGRTGKGLGEAGAP

```

	BASE COUNT	468 a	914 c	902 g	353 t	
	ORIGIN					
	Query Match	4.2%	Score 35;	DB 9;	Length 2637;	
	Best Local Similarity	100.0%;	Pred. No. 1.1e-06;			
	Matches 35; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY		795 ATGTAATCTC	CAAGAAATAA	AAGCAACCAAGAG	829	
Dd		2585 ATGTAATCTC	CAAGAAATAA	AAGCAACCAAGAG	2619	
RESULT 9						
MMCOL18A11						
LOCUS						
DEFINITION			3129 bp DNA linear	ROD 28-MAR-2002		
ACCESSION			Mus musculus alpha-1(XVII) collagen (COL18A1)	gene, exons 40-43,		
VERSION			U03715	complete cds.		
KEYWORDS			U03715.1 GI:4877730			
SEGMENT						
SOURCE			11 of 11			
ORGANISM			Mus musculus			
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;			
AUTHORS			Rehn,M., Hintikka,E. and Pihlajaniemi,T.			
TITLE			Primary structure of the alpha 1 chain of mouse type XVIII collagen, partial structure of the corresponding gene, and comparison of the alpha 1(XVIII) chain with its homologue, the alpha 1(XV) collagen chain			
JOURNAL			J. Biol. Chem. 269 (19), 13929-13935 (1994)			
MEDLINE			94245707			
PUBMED			8188673			
REFERENCE			2 (bases 1 to 3129)			
AUTHORS			Rehn,M., Hintikka,E. and Pihlajaniemi,T.			
TITLE			Characterization of the mouse gene for the alpha 1 chain of type XVIII collagen (Coll18a1) reveals that the three variant N-terminal polypeptide forms are transcribed from two widely separated promoters			
JOURNAL			Genomics 32 (3), 436-446 (1996)			
MEDLINE			96435922			
PUBMED			8838808			
REFERENCE			3 (bases 1 to 3129)			
AUTHORS			Rehn,M.			
TITLE			Direct Submission			
JOURNAL			Submitted (24-NOV-1993) Marko Rehn, University of Oulu, Department of Medical Biochemistry, Kajantieintie 52 A, Oulu, FIN-90220, Finland			
FEATURES			Location/Qualifiers			
source			1. .3129			
			/organism="Mus musculus"			
			/mol_type="genomic DNA"			
			/db_xref="taxon:10090"			
			/clone="X17.6"			
			/tissue type="primary tail culture"			
			/clone_lib="Mouse Cosmid Library in pWE15 by Stratagene (#951303)"			
gene			order(U34606.1:c1346.1755,U34607.2:1..5723, U34608.1:1..2450,U34609.1:1..165,U34610.1:1..3326,			

```
U34612.1:6546. 6572,U34612.1:7451..7486,
U34612.1:7838. 7909,U34612.1:8167..8202,
U34612.1:8463. 8552,U34612.1:9534..9563,
U34612.1:9667. 9693,U34612.1:10015..10077,
U34612.1:11050. 11124,U34612.1:11857..11883,
U34612.1:12083. 12136,U34612.1:12491..12565,
U34612.1:1315. 383,U34613.1:1039..1081,U34613.1:11577..1639,
U34612.1:1253. 1326,U03716.1:1601..1732,U03718.1:262..294,
96..344,942..1139,1653..1768,2515..2884)
/genes="COL18A1"
join(U34607.2:3345..4422,U34608.1:43..587,
U34609.1:28..114,U34610.1:438..491,U34610.1:2478..2607,
U34611.1:490..566,U34611.1:851..1037,U34612.1:227..252,
U34612.1:994..1056,U34612.1:1160..1237,
U34612.1:1461..1514,U34612.1:1671..1829,
U34612.1:2292..2354,U34612.1:3340..3366,
U34612.1:5033..5164,U34612.1:5712..5774,
U34612.1:6546..6572,U34612.1:7451..7486,
U34612.1:7838..7909,U34612.1:8167..8202,
U34612.1:8463..8552,U34612.1:9534..9563,
U34612.1:9667..9693,U34612.1:10015..10077,
U34612.1:11050..11124,U34612.1:11857..11883,
U34612.1:12083..12136,U34612.1:12491..12565,
U34613.1:315..383,U34613.1:1039..1081,U34613.1:11577..1639,
U34612.1:1253..1326,U03716.1:1601..1732,U03718.1:262..294,
96..344,942..1139,1653..1768,2515..2884)
/genes="COL18A1"
join(U34607.2:3705..5163,U34608.1:43..587,
U34609.1:28..114,U34610.1:438..491,U34610.1:2478..2607,
U34611.1:490..566,U34611.1:851..1037,U34612.1:227..252,
U34612.1:994..1056,U34612.1:1160..1237,
U34612.1:1461..1514,U34612.1:1671..1829,
U34612.1:2292..2354,U34612.1:3340..3366,
U34612.1:5033..5164,U34612.1:5712..5774,
U34612.1:6546..6572,U34612.1:7451..7486,
U34612.1:7838..7909,U34612.1:8167..8202,
U34612.1:8463..8552,U34612.1:9534..9563,
U34612.1:9667..9693,U34612.1:10015..10077,
U34612.1:11050..11124,U34612.1:11857..11883,
U34612.1:12083..12136,U34612.1:12491..12565,
U34613.1:315..383,U34613.1:1039..1081,U34613.1:11577..1639,
U34612.1:1253..1326,U03716.1:1601..1732,U03718.1:262..294,
96..344,942..1139,1653..1768,2515..2728)
/genes="COL18A1"
/codon_start=1
/product="alpha-1(XVIII) collagen"
/protein_id="AAC52903.1"
/db_xref="GI:1167906"
/translation="MAPDFERRLCLLLLLSRLVPSADGNSLSPLNPLVWLWPPK
TSDLEGPVQNSPVQSTENPTTHVVDGLTEQTPASSSELPPEEEEDOKA
GQGGSPATVPLVAPASPDMEKNEVAGKILNVAQIRSFVQLWDEDSTIGH
SACTEVPDSGIPTVLPASPOGSKTLLWLSAIPSPDQTEAGTTLAVPTQL
PFPQNLPLGPEPAPPPGPGARFSSSTDGSSWNGOEPPOPHLKGKFLPMTA
RSSQQRHSDVHDHIVPLPLVTGPIVTSLSVHGLLSPSSPSSQLQVQALP
GFGTWSHVAFSGTGLSNDALAGLSLTSRCLPLPPTLLCSRLGIGHFWLPN
HLHTDSVEVETQAWGRFLTNCHPFLAFLLAPSCGPGPPPLPCRCQCEA
LEDCWNLVAGDLPLVLCASLPQOEDGYCVFIIPAAENVAEVLQLLGLDPLPEKIS
QIDPHVGPAYIFGPDNSQVQAYHFKPLFRDRLFLTHVRPATEAGVLPATDA
QVYVSLGVKLSVRDQONLSLTYEGASQCTGTGASFRLPAPFGWTHFALSVDGGS
VALYDCERFQVFPFARASQGLERAGLVQAGTADPDKFGQMSLKVKTGRV
SPVHCLDEDDDDSRASGDFGSESSKHKEDTSLLPLGPQPPVTPPPLAGGSTT
EDPRTETEDDAVDISGATLFTGSSGAWDAIQNGRLIKGMKGKQEPGAGQ
PPGAGPAGVQSGVQSPVQAGPPGQPGKDGTPGRDGPDPGDDGRRG
DTCPGPGTGVGPKGKGDGCIIPGPGPPGPGSPFRDGLTFTDMEGSGSG
DIESLRGPGFGPGPGVGLPGBRFGINGSVAPGAGLPGVGKGGPGPRGP
PIPGPPGKGPVAGQKSGVGDVGIIPGKSGKDLGPIGMEKSKSLAGSPGVQPP
GPPGPGPGPGGAGFGDMEGSGILPWTARSDDGLQGGPPGSLGKDPGVAGLPGA
KGVGADGAGQIPGPGREGAAGSPGKGGKMPGKGNPDKGVGRLPGPGPGG
PVTYVSEDKAIVSTGPGKPGYAGPGPAGPKGDLGSKGEQGLPGKGEKPGTI
FSPDGRALGHPQKAGEPGFPGPGPGYGRPKHGKIGFPGRCRPTNGLKGEKGP
```

CDS

mRNA

```
GDASLGSFMRGLPGPPGPPGPMPIYDSNAFVESGRFOLPQOQGVQSGPKGD
KGEVPPGPPQFPIDLFLHAEMKGDGDRGDAGOKGGERGPGAPGGFFSSVPGP
PGEVPGIPGPKGESIRGPGROGPGGIGYEGROGPGPGPPGPPGPPGPHRQ
TVVPGPPGPPGPPGPGMAGSAGQVIMATYQTMLDKIREVPSGMLIFVAERELY
VVRNGFRKVLLEARTALPGRGNEVAALQPPVLQHGSGPYTRREYSYSTARPWRAD
DILANPPRLPQRPYVPGPHHSSYHLPAPRTLSLAHTHODFQFVLHLVALNTPLS
GSMGRGADQCQCARAVLGSFTFRFLSRLODLSIVRRADRGSPVIYNLKDDEV
LSPGWDLSLFGSQGOLQFQCARIFSDGRDLVRHPAPQKSVMHGSDPSGRLMESYCE
TWRTETTCATGCASSLLSRLLEKQKAAACHNSYIVLCIENSPTSPSK"
Query Match 4.2%; Score 35; DB 10; Length 3129;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 795 ATGTAATCTCTCAAGAAATAAAGGAAGCAAGAG 829
Db 2848 ATGTAATCTCTCAAGAAATAAAGGAAGCAAGAG 2882
RESULT 10
AX409531 3394 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 2178 from Patent WO0229103.
DEFINITION AX409531
ACCESSION AX409531
VERSION AX409531.1 GI:21442236
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2178 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
source
1..3394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. L22548"
BASE COUNT 668 a 1130 c 1080 g 516 t
ORIGIN
Query Match 4.2%; Score 35; DB 6; Length 3394;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 795 ATGTAATCTCTCAAGAAATAAAGGAAGCAAGAG 829
Db 2187 ATGTAATCTCTCAAGAAATAAAGGAAGCAAGAG 2221
RESULT 11
I51045 3394 bp DNA linear PAT 07-OCT-1997
LOCUS Sequence 4 from patent US 5643783.
DEFINITION I51045
ACCESSION I51045
VERSION I51045.1 GI:2472748
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3394)
AUTHORS Olsen,B.R. and Oh,S.P.
TITLE Collagen and uses thereof
JOURNAL Patent: US 5643783-A 4 01-JUL-1997;
FEATURES
source
1..3394
/organism="unknown"
BASE COUNT 668 a 1130 c 1080 g 516 t
ORIGIN
```



```

Query Match      4.2%; Score 35; DB 6; Length 3394;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCTCAAGAAATAAAGGAGCCAAAGAG 829
Db 2187 ATGTAATCTCAAGAAATAAAGGAGCCAAAGAG 2221

RESULT 12
HUMCOL18AX
LOCUS HUMCOL18AX 3394 bp mRNA linear PRI 01-NOV-1994
DEFINITION Human collagen type XVIII alpha 1 (COL18A1) mRNA, partial cds.
ACCESSION L22548
VERSION L22548.1 GI:348908
KEYWORDS alpha-1 type XVIII collagen.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3394)
AUTHORS Oh,S.P., Warman,M.L., Seldin,M.F., Cheng,S.D., Knoll,J.H.,
Timmons,S. and Olsen,B.R.
TITLE Cloning of cDNA and genomic DNA encoding human type XVIII collagen
and localization of the alpha 1 (XVIII) collagen gene to mouse
chromosome 10 and human chromosome 21
JOURNAL Genomics 19 (3), 494-499 (1994)
MEDLINE 94245237
PUBMED 8188291
COMMENT Original source text: Homo sapiens cDNA to mRNA.
FEATURES
Location/Qualifiers
1..3394
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="21q22.3"
/tissue_type="liver"
/dev_stage="adult"
1..3394
/gene="COL18A1"
<1..2055
/gene="COL18A1"
/codon_start=1
/product="collagen type XVIII alpha 1"
/protein_id="AA51864.1"
/db_xref="GI:562794"
/translation="GEVGDGIPGPGPLPGREGIAGPQPKGDRSGRGEKGDPPKXDL
GQPLGPRGPGPVYVSEQGSVLSVPGPRRGFGFPAGPKGNLGSKEIGS
PGPKGEKPGSIFSPDQALGPAQKAGKEFGFRGPGLYGRPGYKGEIGFPPGR
PGMGLKGEKSGDASLGFGRMGPMGPPGPPGPGTPTVYDNNVFAESRPPGPG
LPGNQGPPGPKGKEGVPGPQGFPPDFLOKEAMKGEKGDGADGKGRGPPG
GGFFSSLPGAPGAPRGYCGI PGPKGESIRGQPGPGPGPGIGYEGRQGP
GPPGPPFPGRQITISVPGPGPGPGPGPGPGTGMAGSSQVRLMATRQAMLQVHEVP
EGWLLFVAEESLYRVQNGFRKQLEARTPLRGTDNEAALQPPVQLHDSNPYPR
REHPHTAPRWADDIASPPGPFPGPHHSYVHCGPARTSPPAHSRDFQ
PVLHLVALNSLGGMRGTGADFCFOQARAVGLAGTFRFLSRLQDLIVRRAD
RAAPVNLKDLLPFWELFSGSEGLPKGARI FSPDKDLRHPTWPKSVWHGS
DNGRLTESYCTWRTAPSATQASSLLGRLGQSAASCHHAYIVLCIENSFWTA
SK"
BASE COUNT 668 a 1130 c 1080 g 516 t
ORIGIN

Query Match      4.2%; Score 35; DB 9; Length 3394;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCTCAAGAAATAAAGGAGCCAAAGAG 829
Db 2187 ATGTAATCTCAAGAAATAAAGGAGCCAAAGAG 2221

RESULT 13
HUMCOL18AX
LOCUS HUMCOL18AX 3394 bp mRNA linear PRI 01-NOV-1994
DEFINITION Human collagen type XVIII alpha 1 (COL18A1) mRNA, partial cds.
ACCESSION L22548
VERSION L22548.1 GI:348908
KEYWORDS alpha-1 type XVIII collagen.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3394)
AUTHORS Oh,S.P., Warman,M.L., Seldin,M.F., Cheng,S.D., Knoll,J.H.,
Timmons,S. and Olsen,B.R.
TITLE Cloning of cDNA and genomic DNA encoding human type XVIII collagen
and localization of the alpha 1 (XVIII) collagen gene to mouse
chromosome 10 and human chromosome 21
JOURNAL Genomics 19 (3), 494-499 (1994)
MEDLINE 94245237
PUBMED 8188291
COMMENT Original source text: Homo sapiens cDNA to mRNA.
FEATURES
Location/Qualifiers
1..3394
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="21q22.3"
/tissue_type="liver"
/dev_stage="adult"
1..3394
/gene="COL18A1"
<1..2055
/gene="COL18A1"
/codon_start=1
/product="collagen type XVIII alpha 1"
/protein_id="AA51864.1"
/db_xref="GI:562794"
/translation="GEVGDGIPGPGPLPGREGIAGPQPKGDRSGRGEKGDPPKXDL
GQPLGPRGPGPVYVSEQGSVLSVPGPRRGFGFPAGPKGNLGSKEIGS
PGPKGEKPGSIFSPDQALGPAQKAGKEFGFRGPGLYGRPGYKGEIGFPPGR
PGMGLKGEKSGDASLGFGRMGPMGPPGPPGPGTPTVYDNNVFAESRPPGPG
LPGNQGPPGPKGKEGVPGPQGFPPDFLOKEAMKGEKGDGADGKGRGPPG
GGFFSSLPGAPGAPRGYCGI PGPKGESIRGQPGPGPGPGIGYEGRQGP
GPPGPPFPGRQITISVPGPGPGPGPGPGPGTGMAGSSQVRLMATRQAMLQVHEVP
EGWLLFVAEESLYRVQNGFRKQLEARTPLRGTDNEAALQPPVQLHDSNPYPR
REHPHTAPRWADDIASPPGPFPGPHHSYVHCGPARTSPPAHSRDFQ
PVLHLVALNSLGGMRGTGADFCFOQARAVGLAGTFRFLSRLQDLIVRRAD
RAAPVNLKDLLPFWELFSGSEGLPKGARI FSPDKDLRHPTWPKSVWHGS
DNGRLTESYCTWRTAPSATQASSLLGRLGQSAASCHHAYIVLCIENSFWTA
SK"
BASE COUNT 668 a 1130 c 1080 g 516 t
ORIGIN

Query Match      4.2%; Score 35; DB 6; Length 4031;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCTCAAGAAATAAAGGAGCCAAAGAG 829
Db 2187 ATGTAATCTCAAGAAATAAAGGAGCCAAAGAG 4021

RESULT 14
MUSCOLPRO
LOCUS MUSCOLPRO 4031 bp mRNA linear ROD 18-JUL-1994
DEFINITION Mouse alpha 1 (XVIII) collagen mRNA, partial cds.
ACCESSION L22545
VERSION L22545.1 GI:348968
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 4031)
AUTHORS Oh,S.P., Kamagata,Y., Muragaki,Y., Timmons,S., Ooshima,A. and
Olsen,B.R.
TITLE Isolation and sequencing of cDNAs for proteins with multiple
domains of Gly-Xaa-Yaa repeats identify a distinct family of
collagenous proteins
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91 (10), 4229-4233 (1994)
MEDLINE 94240111
PUBMED 8183893
COMMENT Original source text: Mus musculus cDNA to mRNA.
FEATURES
Location/Qualifiers
1..4031
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="liver"
/dev_stage="embryo"
<1..3867
/codon_start=1
/product="alpha 1 (XVIII) collagen"
/protein_id="AAA13787.1"
/db_xref="GI:511298"
/translation="ENVAEEVGLLQLGDLPLPEKISQIDDPHVGPAYIFGPDNSGQV
AQYHPKLFRRDFSLFHVPRATEAAGVLFAITDAQVGVSLGKLVSRDQGNISL
LYTEPGASOTOTGASFRLPAPVFGOWTFALSDVGGSVLYVDCERFORVPFARASQGL
ELERAGLIVGQAGTADPKQGMISELKVKTFRVSPVHCLDEDDDEDASGDFGS
FGESSKHKEDTSLPLGPFPVTPSPPLAGSTTDPRTETSEDAVDSIGATLL
PFGSSGAWDEAIQNPGRGLKMGKQKGEAGQPGPAGPQGPAGPVVGSFNSQPL
VPAQPGPGPGPGKGTGPDGPDGPDGPDGPDGPDGPDGPDGPDGPDGPDGPDGPDG
PGIGRPGPGPGPGPGPGPGPKLTFIDMGSGFSGDIESLRGPGFPGPGPGPGVPG
LPGEPRGINGSYAPGAGLPGVPGKSGAGPGPGPGPGPGPGPGPGPGPGPGPGAGDMEG
GDVIGIPKSGKGLDLPICMGKSGSLAGSPGVPGPGPGPGPGPGPGPGPGPGPGPGAGDMEG
SGIPLWTTARSSDGLQGGPPGSLKGDGVAGLPGKAGVEADGAGGIPGPGREGAA
GSPGPKGKGMFGKNGKGVGRLPGPGLPVIYSSSEKAIYSTPGPEKPK
GYAGPPGPGPKGDLGSGKEQGLPGFKGKEPGTIFSPDGRRLGHPKQKAGKEGPR

```

GPPGVPGRGPKHKEIGFPCRGPRGTNGLKGKGRBPASLGF5MRGLPGPPGPPGPP
 GPGWPIYDSNAFVSGRGLGQOQVQSPGKDGKGVGPPGPPGPPIDLEHLEA
 ENKDGKGRGDAGQKGERCEPCAGGFFSSSVPPGPPGPGYPIPGKGESIRGPPG
 PPGPPGPIGVEGQKGGPPGPPGPPGPHRQTVSVPPGPPGPPGPPGPPGPPGPPG
 SAGVRIWATYQTMLDKREVPEGLIFVAREEELVVRNGFRKVLLEARTALLRG
 GNEVAAFQPLVQLHKGESPYTRREYSYSTARPRADDILANPRILPDRQPYGVPVPHH
 SSVHLPPARPTLSLAHTHODFQPVHLVALNTPLSGGMRGIRGADFOCQOARAVGL
 SCTFRFLSSRLQDLYSIVRRADRGSPVIVNLKDEVLSPSWDSLFGSQGOVQPCARI
 FSDGRDVRHPAWOKSVVHSGSDSGRLMESYCTETWRTTGTGATGQASSLSRLLL
 EOKAASCHNSYIVLCIENSFMFSK"

BASE COUNT 873 a 1167 c 1268 g 723 t
 ORIGIN

Query Match 4.2%; Score 35; DB 10; Length 4031;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCTCAAGAAATAAAGCAAGCAAGAG 829

DB 3987 ATGTAATCTCAAGAAATAAAGCAAGCAAGAG 4021

RESULT 15

MUSCOLLAG 4437 bp mRNA linear ROD 04-FEB-2003
 LOCUS Mus musculus mRNA for collagen, partial cds.
 DEFINITION D17546
 ACCESSION D17546.1 GI:2160436

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 4437)

ABE.N., Muragaki, Y., Yoshioka, H., Inoue, H. and Ninomiya, Y.

Identification of a novel collagen chain represented by extensive

interruptions in the triple-helical region

Unpublished

2 (bases 1 to 4437)

ABE.N.

Direct Submission

Submitted (04-SEP-1993) Nobuhiro Abe, Okayama University Medical

School, Molecular Biology and Biochemistry; 2-5-1, Shikata-cho,

Okayama, Okayama 700, Japan (Tel:81-86-223-7151(ex.2390),

Fax:81-86-222-7768)

On Jun 5, 1997 this sequence version replaced gi:467516.

COMMENT

FEATURES

source

1. .4437
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /dev_stage="embryo"
 /notes="clones NA[1.12.286.261]"
 <1. .3423

CDS

/codon_start=1
 /product="collagen"
 /protein_id="BAA04483.1"
 /db_xref="GI:467517"
 /translation="QRVFFARASQGLEBAGLVFGQAGTADPDKFCQMI BELKVRK
 TPVSPVHCDDEDDSDASGDFSGFESSKSHKEDTSLPGLPQPPVTPSPPLAG
 GSTTEPRTEEDAAVDSIGAEITLPGTSGAWDEAIQNPGRGLIKGMKGKGE
 GAQGPAGPAGPAGPVQSPNSQVPGNAQPPGPGQKGTGPRDGEPCDGED
 GRGDTGPGQFTGPDGVGKKGKDPGIGRPGPPGPPGPPGPPGPPGPPGPPGPPG
 GFGSDLESGLRFPFGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
 FPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
 VGPFGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
 LPGAKEVAGDAAGQIPGPPREGAAGSPGKKGKGNKGNKGNKGNKGNKGNKGNKGNK
 GPBGVTVYSSDKAIVTSPGPKPCYAGPCGPGKGLGSKGQGLGPKGKGE
 PGTFPDRALGHPQKAGKGEPRGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
 KGEFGASLGRMRGLPFPFGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
 PKDGKGVGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
 VPFGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
 PHRTQTVSPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
 PHRTQTVSPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG

BELVVRVNRGPKVLEARTALPRGTGNEVAALOPPLVOLHEGSPYTRREYSYSTARP
 WRADDILANPRPTLDRQPYGVPVPHHSSVHLPPARPTLSLAHTHODFQPVHLVALN
 TPLSGMRGIRGADFOCQOARAVGLSGTFRFLSSRLQDLYSIVRRADRGSPVIVNL
 KDEVLSPSWDSLFGSQGOQLOPGARIFSDGRDVRHPAWOKSVVHSGSDSGRLME
 SYCTETWRTTGTGATGQASSLSRLLEQKAAASCHNSYIVLCIENSFMFSK"

3'UTR 3424. .4437

polyA_signal /evidence=experimental

polyA_signal 3558. 3563

polyA_signal /evidence=experimental

polyA_signal 4415. .4420

polyA_signal /evidence=experimental

polyA_signal 4437

polyA_signal /evidence=experimental

BASE COUNT 962 a 1311 c 1345 g 819 t

ORIGIN

Query Match 4.2%; Score 35; DB 10; Length 4437;

Best Local Similarity 100.0%; Pred. No. 1.1e-06;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCTCAAGAAATAAAGCAAGCAAGAG 829

DB 3543 ATGTAATCTCAAGAAATAAAGCAAGCAAGAG 3577

Search completed: August 17, 2003, 18:39:17

Job time : 4164.77 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 12:12:16 ; Search time 319.261 Seconds
(without alignments)
7009.423 Million cell updates/sec

Title: US-09-938-391-1

Perfect score: 829

Sequence: 1 cctcgccggcgagatgacat.....aataaagaagccaaagag 829

Scoring table: IDENTITY NUC

Gapop 10.0., Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
- 25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	829	100.0	829	24	Canine pro-endosta
2	555	66.9	555	24	Canine endostatin
3	550.4	66.4	552	21	Canine angiogenesi
4	503.2	60.7	4875	24	Human ovarian anti
5	503.2	60.7	5408	20	Human collagen 18
6	496.8	59.9	3394	18	Human alpha-1 coll
7	496.8	59.9	3394	20	Human alpha1 (XVII
8	496.8	59.9	3394	24	Gene #2178 used to

9	493.2	59.5	900	24	ABK09977	Synthetic plasmid
10	493.2	59.5	4551	24	ABV94763	Human pancreatic c
11	493.2	59.5	4551	24	ABN85301	Human collagen XVI
12	425.6	51.3	968	21	ABQ62025	Nucleotide sequenc
13	423.6	51.1	1564	24	ABQ76740	DNA encoding human
14	423	51.0	4031	18	AA184485	Mouse alpha-1 coll
15	421.2	50.8	816	21	AAA64013	DNA encoding a hum
16	420	50.7	555	22	AA166529	Human vascular end
17	419.2	50.6	558	24	ABA99261	Human endostatin c
18	418.4	50.5	563	20	AAZ08750	Human endostatin c
19	418	50.4	546	21	AAA27004	Human endostatin g
20	418	50.4	549	20	AAZ77719	Human endostatin D
21	418	50.4	549	21	AAA29884	Human angiogenesis
22	418	50.4	549	21	AAZ51291	Human angiogenesis
23	418	50.4	549	22	AAZ00867	Human gene fragmen
24	418	50.4	549	25	ABA00774	Human endostatin c
25	418	50.4	552	20	AAZ35375	SEQ ID 50 of WO991
26	418	50.4	552	21	AAZ62023	Nucleotide sequenc
27	418	50.4	552	21	AAZ68203	Human endostatin c
28	418	50.4	552	22	AAZ88289	Human endostatin c
29	418	50.4	552	24	ABK50685	cDNA encoding huma
30	417.2	50.3	641	22	AAH79104	Human endostatin e
31	417.2	50.3	641	24	AAZ5454	Specific tumour ce
32	417.2	50.3	641	24	AAZ44000	Reproductive recom
33	417.2	50.3	641	24	ABQ76079	Anticancer gene-as
34	417	50.3	551	24	ABQ81193	Human endostatin c
35	412.2	49.7	6462	21	AAA64014	Nucleotide sequenc
36	410.8	49.6	534	21	AAZ27005	Alternate human en
37	410.8	49.6	537	22	AAZ00868	Human gene fragmen
38	398.4	48.1	650	24	ABK09978	Synthetic plasmid
39	394.2	47.6	552	21	AAZ50398	Human endostatin c
40	376.8	45.5	573	21	AAZ67777	Murine endostatin
41	376.8	45.5	573	24	ABK47719	cDNA encoding mous
42	376.6	45.4	565	20	AAZ58740	DNA encoding anti-
43	376.4	45.4	580	20	AAZ08748	Mouse endostatin c
44	376.2	45.4	558	22	AAZ18701	Mouse endostatin e
45	376.2	45.4	558	24	ABQ76173	Rat endostatin enc

ALIGNMENTS

RESULT 1	
AAZ46062	
ID	AAZ46062 standard; cDNA; 829 BP.
XX	AAZ46062;
AC	AAZ46062;
XX	19-JUL-2002 (first entry)
DT	Canine pro-endostatin coding sequence.
XX	
DE	
XX	
KW	Dog; pro-endostatin; endostatin; angiogenesis; cancer; gene therapy;
KW	psoriasis; rheumatoid arthritis; retinopathy; macular degeneration;
KW	corneal graft rejection; neovascular glaucoma; retrolental fibroplasia;
KW	rubeosis; Osler-Weber Syndrome; myocardial angiogenesis;
KW	plaque neovascularisation; telangiectasia; haemophilic joints;
KW	angiofibroma; wound granulation; coronary collateral;
KW	cerebral collateral; arteriovenous malformation;
KW	ischemic limb angiogenesis; diabetic neovascularisation; fracture;
KW	cytostatic; antiarthritic; antipsoriatic; antidiabetic; ophthalmological;
XX	gynaecological; gene; ss.
XX	
OS	Canis familiaris.
XX	
Key	Location/Qualifiers
CDS	1..693
FT	/*tag= a
FT	/product= "pro-endostatin"
FT	/partial
FT	/note= "no start codon"
XX	
PN	EP1191036-A2.


```
FT      /*tag= a
FT      /product= "endostatin/IgG1Fc construct"
FT      /transl_except= (pos:58..60,aa:Ala)
FT      /transl_except= (pos:811..816,aa:GGGSGG)
XX
XX      CN1354186-A.
XX      PN
XX      PD
XX      19-JUN-2002.
XX      30-NOV-2000; 2000CN-01233347.
XX      30-NOV-2000; 2000CN-01233347.
XX      (LIAO-) LIAONING WEIXING BIOLOGICAL PROD INST CO.
XX      Chen L, Li Z, Liu Q;
XX      WIPI; 2002-751441/82.
XX      P-PSDB; ABG73586.
XX      Preparation of recombinant human vascular endothelial cytopoiesis
XX      suppressor factor with human IgG1Fc fragment molecular structure and
XX      application of its product -
XX      Disclosure; Page 7-8 (Disclosure); 12pp; Chinese.
XX
XX      This invention describes a novel method for the preparation of
XX      recombinant human vascular endothelial cytopoiesis inhibiting factor with
XX      human IgG1Fc fragment molecular structure and its product application.
XX      The novel factor is derived from endostatin (using PCR to screen a human
XX      foetal kidney cell cDNA library) and human IgG1Fc. The product of the
XX      invention can specifically inhibit tumour vascular endothelial
XX      proliferation and can be used for curing several tumour types. This
XX      sequence encodes a fusion construct composed of human endostatin and
XX      human IgG1Fc, described in the disclosure of the invention.
XX
XX      Sequence 1564 BP; 346 A; 491 C; 436 G; 291 T; 0 other;
XX
XX      Query Match      51.18; Score 423.6; DB 24; Length 1564;
XX      Best Local Similarity 77.9%; Pred. No. 4.2e-61;
XX      Matches 528; Conservative 0; Mismatches 139; Indels 11; Gaps 1;
XX
XX      137 CCACACCCACACAGGACTTCCAGTGGTGTGCACCTGTGGCCCTGAACAGCCGCGAGC 196
XX      260 CTACAGCCACCGGACTTCCAGCCGGTGTCTCCACCTGGTGGCTCAACAGCCCTGT 319
XX
XX      197 CGGCGCGCATCGAGGATCCGGGAGCGGACTTCCAGTGTCCAGTGTCCAGCGCGCGCG 256
XX      320 CAGCGCGCATCGGGGATCCGGGGGCGGACTTCCAGTGTCTCCAGCAGCGCGGGCGCG 379
XX
XX      257 CGGGCTGGCGGACACTTCCGGGCTTCTGTGCTCGGCTGCGAGCTCTACAGCA 316
XX      380 TGGGGCTGGCGGACACTTCCGGGCTTCTGTGCTCGGCTGCGAGCTCTACAGCA 439
XX
XX      317 TCGTGGCGCGCGGCGGACCGGCGGCTGCTGTCAACCTCAGGACGAGTGTCT 376
XX      440 TCGTGGCGCGGCGGACCGGCGGCTGCTGTCAACCTCAGGACGAGTGTCT 499
XX
XX      377 TCCCCAGCTGGAGGCTTATTTCTGGGCTTCGAGGCGCCAGCTGAAGCCCGGCGCGCA 436
XX      500 TTCCAGCTGGAGGCTCTGTCTTCTCAGGCTCTGAGGCTCGGCTGAAGCCCGGCGCGCA 559
XX
XX      437 TCTTCTTTTCGAGGCGAGATGCTCTGCGAGCACCCCGCTGCGCCCGGAGAGCGTGT 496
XX      560 TCTTCTCTTTGAGCGGCAAGAGCGTCTGAGGCGACCCCACTTGGCCCCCAGAGAGCGTGT 619
XX
XX      497 GGCACGCGCTCGAGCCCGGCGGCGGCGGCTGAGCCGACAGCTACTGCGAGACGCTGGCGGA 556
XX      620 GGCATGCTCGAGCCCGGCGGCGGCGGCTGAGCCGACAGCTACTGTCGAGACGCTGGCGGA 679
XX
XX      557 CGAGAGCCCGGCGGCGGCGGCGGCGGCTGCTGCTGCTGCGGCGGCGGCGGCTGCGAGC 616
XX      680 CGAGAGGCTCGGCGGCGGCGGCGGCGGCGGCTGCTGCTGCTGCGGCGGCGGCGGCTGCTGCGGCG 739
```

```
QY      617 AGGAGSCCGGAGCTGCCGCCAGCGCTTCTGTGTCTGCTGATCGAGAACAGCGTTCATGA 676
DB      740 AGAGTGCCCGGAGCTGCCATCAGCTATCGTGTCTGTGATTGAGAACAGCTTCATGA 799
QY      677 CTTCTCTTCTCCAGTAGGCGCGCGGCCACGAGCAGCGGGGAGGGCGCCCGCAG 736
DB      800 CTGCTCTCCAGTAGTAAGG-----AGGAAGCGGAGGTGGAGGGTCCGCTCGACAA 848
QY      737 GAGCATCCCGCGCGCGGGGGCGCTGGCGGGGAGCGCTTGCCTGCGACCGTTCAGTTTAT 796
DB      849 AACTCACATGCGCCACCGTGCACACCTGAACTCTCTGGGGGAGCCGCTCAGTCTTCT 908
QY      797 GTAATCTCTCAAGAAATAA 814
DB      909 CTTCCCGCCCAAAACCCAA 926
```

RESULT 14

AAT84485

ID AAT84485 standard; cDNA; 4031 BP.

AC AAT84485;

XX 19-NOV-1997 (first entry)

XX Mouse alpha-1 collagen cDNA.

DE Alpha-1 collagen; type XVIII collagen; cartilage degeneration; ss.

XX Mus musculus.

XX Key Location/Qualifiers

FT CDS 1..3867

FT /*tag= a

XX US5643783-A.

XX 01-JUL-1997.

XX 01-DEC-1993; 93US-0159784.

XX 01-DEC-1993; 93US-0159784.

XX (HARD) HARVARD COLLEGE.

XX Oh SP, Olsen BR;

XX WIPI; 1997-350247/32.

XX P-PSDB; AAW26328.

XX Nucleic acid encoding human alpha-1 collagen - for production of

XX recombinant alpha-1 collagen, for use in the treatment of cartilage

XX degeneration

XX Disclosure; Fig 2; 35pp; English.

XX A cDNA clone (AAT84485) codes for a mouse novel type alpha-1 (XVIII)

XX collagen (AAW26328) that is expressed in multiple organs, especially

XX the liver, lung and kidney. It was isolated by screening murine

XX 15.5- and 17.5-day embryo libraries with probes based on murine

XX type XII collagen and on human alpha-1 collagen cDNA. Isolated

XX nucleic acids can be used to express recombinant alpha-1 collagen

XX in transformed host cells. Claimed nucleic acid (see AAT84484)

XX coding for human alpha-1 collagen (AAW26327) can be used to treat

XX patients suffering from diseases associated with degradation of

XX cartilage, and for supplementing collagen.

XX Sequence 4031 BP; 873 A; 1172 C; 1266 G; 720 T; 0 other;

XX Query Match 51.0%; Score 423; DB 18; Length 4031;

XX Best Local Similarity 76.9%; Pred. No. 4.9e-61;

XX Matches 548; Conservative 0; Mismatches 150; Indels 15; Gaps 2;

QY 1 CCTGGCGGCGAGATGACATCTCTGGCGGCGGCCCCCGGCTGTGACCCCGAGCCCTAC 60
Db 3184 CCTGGCGGCGAGATGACATCTCTGGCGGCGGCCCCCGGCTGTGACCCCGAGCCCTAC 3243
QY 61 CCGGGG-----CCCGGACACAGGCTCTAGTGTGACTTCAGCGGCTGCCCCACTGGT 117
Db 3244 CTTGGAGTTCCACATCACCACAGTTCTTATGTGACCTGCGCGCAGCCCGCCCCAC- --- 3299
QY 118 GGGCCCGTCCACACCCACACCCACCCAGGACTTCAGTGTGTGACGCTGGTG 177
Db 3300 -----CCTCTCATCTGCTATATCATCAGGACTTCAGCGAGTGTCCACCTGGTG 3351
QY 178 GCCTTGAACAGCCCGGCGGCGGCGGCTGCGAGGATCCGGGAGCGGACTTCAGTGC 237
Db 3352 GCACTGAACACACCCCTGTCTGGAGCATGCGTGTATCGTGGAGCAGATTTCAGTGC 3411
QY 238 TTCCACAGGCGGCGGCGGCGGCTGCGGCGGACTTCGGGCTTCTGTGTCGCGG 297
Db 3412 TTCCACAGGCGGCGGCGGCTGCGGCGGACTTCGGGCTTCTGTGTCGCTAGG 3471
QY 298 CTGACGAGCTCTACAGATCTGTGCGCGCGCGACCGACCGGGGTGCGCTGCAAC 357
Db 3472 CTGACGATCTCTATAGATCTGTGCGCGGCTGTGACCGGGGTCTGTGCCATCTGCAAC 3531
QY 358 CTGAGGAGAGGTGCTTCTCCAGCTGGGAGGCTTATTTCTCGGGCTCCGAGGCCAG 417
Db 3532 CTGAGGAGAGGTGCTTCTCCAGCTGGGACTCTCTTTCTGGCTCCAGGTCGAA 3591
QY 418 CTGAAGCGGCGGCGGCGGCTTCTTCTTTCGAGCGGAGATGCTCTGAGCACCCCGC 477
Db 3592 GTGCAACCGGCGGCGGCGGCTTCTTCTTTCGAGCGGAGATGCTCTGAGCACCCAGC 3651
QY 478 TGSCCCCGGAAGCGGTGTGCGACGCTCGACCCCGCGGCGCGCTGACCGACAGC 537
Db 3652 TGSCCCCGGAAGCGGTGTGCGACGCTCGACCCCGCGGCGCGCTGAGGAGT 3711
QY 538 TACTGCGAGAGCTGTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 597
Db 3712 TACTGTGAGACATGCGGAACTACTGTGGGCTACAGGTACAGGCTCTCTGCTG 3771
QY 598 GCGGCGAGGCTGTGAGAGAGGCGGCGGAGTGTGCGGCGGCGGCGGCGGCGGCGGCG 657
Db 3772 TCAGGAGGCTCTGGAACAGAAAGTGTGAGTGTGCGGCGGCGGCGGCGGCGGCGG 3831
QY 658 ATCGAAGACAGCTGATGCTCTCTTCTCAAGTAGGCGGCGGCGGCGGCGGCGGCGG 710
Db 3832 ATTGAGATAGCTTCATGACCTCTTTCTCAATAGGCTCTGCGGAGTAGGG 3884

RESULT 15
AAA64013
ID AAA64013 standard; DNA; 816 BP.
XX AC AAA64013;
XX
XX
XX 20-DEC-2000 (first entry)
DE DNA encoding a human angiogenesis inhibitor designated KED.
XX
KW Angiogenesis inhibitor; KED; kringle protein; angiogenesis; tPA;
KW tissue plasminogen activator; tumour; atherosclerosis; arthritis;
KW retinopathy; bronchial vascular congestion; inflammatory bowel disease;
KW adult respiratory distress syndrome; Castlemans disease; psoriasis;
KW hepatitis; aneurysm; renal disease; haemangioma; ss.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX 1..816
XX FT CDS
XX FT /*tag= a
XX FT /product= "angiogenesis inhibitor KED"

PN WO20049871-A1.
XX 31-AUG-2000.
XX
XX 24-FEB-2000; 2000WO-US04798.
XX
XX 24-FEB-1999; 99US-0121341.
PR 25-FEB-1999; 99US-0121633.
PR 18-NOV-1999; 99US-0166176.
XX
XX (FORD-) FORD HEALTH SYSTEM HENRY.
XX
PI Dou D, Chopp M, Wang L, Mikkelsen T;
XX WPI; 2000-572016/53.
DR P-PSDB; AAB08407.
XX
XX Use of kringle protein and kringle derived from plasminogen and
PT composition comprising kringle proteins for treating tumor and
PT atherosclerosis, arthritis and retinopathy -
XX
PS Disclosure; Page 151; 163pp; English.
XX
CC The present sequence encodes a human polypeptide which is a potent
CC angiogenesis inhibitor, designated KED. The KED polypeptide, kringle
CC proteins, or a kringle derived from human tissue plasminogen activator
CC (tPA) protein are used to inhibit angiogenesis. Kringle protein, KED
CC or tPA are useful for treating tumours, as well as atherosclerosis,
CC arthritis, retinopathy and other similar diseases. KED is also useful
CC for the treatment of diseases such as bronchial vascular congestion,
CC inflammatory bowel disease, adult respiratory distress syndrome,
CC Castlemans disease, psoriasis, hepatitis, aneurysm, renal disease
CC and haemangioma.
XX
SQ Sequence 816 BP; 156 A; 259 C; 254 G; 147 T; 0 other;
Query Match 50.8%; Score 421.2; DB 21; Length 816;
Best Local Similarity 85.0%; Pred. No. 1.1e-60;
Matches 471; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
QY 131 CCCACACCACACCCACAGGACTTCCAGTGTGCTGTGACCTGTGGGCTTGAACAGCC 190
Db 257 CCCATATGACAGCCACCGGACTTCCAGCGGTGTCTCCACCTGTGGCTCAACAGCC 316
QY 191 CGCAGCGCGGCGGCTGCGAGGATCCCGGGAGCGGACTTCCAGTGTTCAGCAGCGC 250
Db 317 CCTGTGAGCGGCTGCGGCGGATCCCGGGCGGAGTCCAGTGTTCAGCAGCGC 376
QY 251 GCGCGCGGCGGCTGCGGCGGACTTCCCGGCGCTTCTGTGCTGCGGCTGCGAGGCTCT 310
Db 377 GGGCGGTGGGCTGCGGCGGACTTCCCGGCGCTTCTGTGCTGCGGCTGCGAGGCTGT 436
QY 311 ACAGATGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 370
Db 437 ACAGCATGTGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 496
QY 371 TGTCTCTTCCAGCTGGGAGGCTTATTCTCGGCGCTCCGAGGCGGCGGCGGCGGCGG 430
Db 497 TGTCTTCTCCAGCTGGGAGGCTTGTCTTCTAGGCTCTGAGGCTCGCTGAGGCGGCGG 556
QY 431 CCGCATCTTCTCTTTCGACGCGGAGATGTCTGTGAGCACCCCGGCTGCGGCGGCGGCGG 490
Db 557 CAGCATCTTCTCTTTCGACGCGGAGGAGTCTCTGAGGCGGCGGCGGCGGCGGCGGCGG 616
QY 491 GGTGTGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 550
Db 617 GGTGTGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 676
QY 551 GCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 610
Db 677 GCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 736
QY 611 TGGAGCAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 670

Db 737 TGGGGCAGAGTGGCGGAGCTGCCATCACGCCCTACATCGTGTCTCTGCATTGGAACAGCT 796
QY 671 TCATGACCTCCTTC 684
Db 797 TCATGACTGCCTCC 810

Search completed: August 17, 2003, 13:19:25
Job time : 324.261 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 14:11:23 ; Search time 319.86 Seconds
(without alignments)
6996.296 Million cell updates/sec

Title: US-09-938-391-1

Perfect score: 829

Sequence: 1 cccctggcggcagatgacat.....aataaaaggacccaagag 829

Scoring table: OLIGO.NUC

Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	829	100.0	829	24 AAL46062	Canine pro-endostatin
2	555	66.9	555	24 AAL46063	Canine endostatin
3	529	63.8	552	21 AA251309	Canine angiogenesis
C 4	35	4.2	35	24 AAL46065	Murine collagen XV
C 5	35	4.2	251	24 ABN94215	Gene #713 used to
C 6	35	4.2	338	22 AAK53858	Murine transport a
7	35	4.2	697	24 ABQ55585	Human ovarian anti
8	35	4.2	3394	18 AAT84484	Human alpha-1 coll

9	35	4.2	3394	20 AAX78379	Human alpha (XVII
10	35	4.2	3394	24 ABN95680	Gene #2178 used to
11	35	4.2	4031	18 AAT84485	Mouse alpha-1 coll
12	35	4.2	4875	24 ABQ54955	Human ovarian anti
13	35	4.2	5408	20 AAX77720	Human collagen 18
C 14	35	4.2	37664	22 AAK83781	Human immune/haema
C 15	35	4.2	61710	22 AAK83782	Human immune/haema
16	29	3.5	552	21 AA250398	Murine endostatin c
17	27	3.3	27	24 AAL46064	Murine collagen XV
18	27	3.3	434	24 ABV94631	Human pancreatic c
19	27	3.3	574	24 ABV94691	Human pancreatic c
20	27	3.3	900	24 ABK09977	Synthetic plasmid
21	27	3.3	4551	24 ABV94763	Human pancreatic c
22	27	3.3	4551	24 ABN85301	Human collagen XVI
23	26	3.1	534	21 AAA27005	Alternate human en
24	26	3.1	537	22 AAS00868	Human gene fragmen
25	26	3.1	549	22 AAS00867	Human gene fragmen
26	26	3.1	549	25 ABA00774	Human endostatin c
27	26	3.1	546	21 AAX27004	Human endostatin g
28	26	3.1	549	21 AAA29884	Human angiogenesis
29	26	3.1	549	21 AAZ51291	Human angiogenesis
30	26	3.1	549	22 AAS00867	Human gene fragmen
31	26	3.1	551	24 ABQ81193	Human endostatin c
32	26	3.1	552	20 AAX35375	SEQ ID 50 of WO991
33	26	3.1	552	21 AAC62023	Nucleotide sequenc
34	26	3.1	552	21 AAA68203	Human endostatin c
35	26	3.1	552	22 AAC88289	Human endostatin c
36	26	3.1	552	24 ABK50685	cDNA encoding huma
37	26	3.1	555	22 AAI66529	Human vascular end
38	26	3.1	558	24 ABA99261	Human endostatin c
39	26	3.1	563	20 AAZ08750	Human endostatin c
40	26	3.1	641	22 AAH79104	Human endostatin c
41	26	3.1	641	24 AAL55454	Specific tumour ce
42	26	3.1	641	24 AAL44000	Reproductive recom
43	26	3.1	641	24 ABQ76079	Anticancer gene-as
44	26	3.1	650	24 ABK09978	Synthetic plasmid
45	26	3.1	816	21 AAA64013	DNA encoding a hum

ALIGNMENTS

RESULT 1
AAL46062
ID AAL46062 standard; cDNA; 829 BP.

XX AAL46062;

XX 19-JUL-2002 (first entry)

XX Canine pro-endostatin coding sequence.

XX Dog; pro-endostatin; endostatin; angiogenesis; cancer; gene therapy;
KW psoriasis; rheumatoid arthritis; retinopathy; macular degeneration;
KW corneal graft rejection; neovascular glaucoma; retrolental fibroplasia;
KW rubeosis; Osler-Webber Syndrome; myocardial angiogenesis;
KW plaque neovascularisation; telangiectasia; haemophilic joints;
KW angiofibroma; wound granulation; coronary collateral;
KW cerebral collateral; arteriovenous malformation;
KW ischaemic limb angiogenesis; diabetic neovascularisation; fracture;
KW cytostatic; antiarthritic; antipsoriatic; antidiabetic; ophthalmological;
KW gynaecological; gene; ss.

OS Canis familiaris.

XX Location/Qualifiers
Key 1..693
CDS /*tag= a
/product= "pro-endostatin"
/partial
/note= "no start codon"

PN EP1191036-A2.

QY 342 GGTGCGCGTCTCAACTCAGGACGAGGTCTCTTCCAGCTGGAGGCTTATTTCTC 401
DB |||||||
QY 204 GGTGCGCGTCTCAACTCAGGACGAGGTCTCTTCCAGCTGGAGGCTTATTTCTC 263
DB |||||||
QY 402 GGGCTCCGAGGGCCAGCTGAAGCCCGGGGCCCGCATCTTCTTTTCGACGGCAGAGATGT 461
DB |||||||
QY 264 GGGCTCCGAGGGCCAGCTGAAGCCCGGGGCCCGCATCTTCTTTTCGACGGCAGAGATGT 323
DB |||||||
QY 462 CTTGACGACCCCGCTTGGCCCGGAGAGCGTGTGGCAGCGCTTCGACCCCGAGCGGGCG 521
DB |||||||
QY 324 CTTGACGACCCCGCTTGGCCCGGAGAGCGTGTGGCAGCGCTTCGACCCCGAGCGGGCG 383
DB |||||||
QY 522 CCGCTTACGACGAGCTTACTGCGAGAGCTGCGGAGCGGAGCGGCGGCGGCGGCGGCA 581
DB |||||||
QY 384 CCGCTTACGACGAGCTTACTGCGAGAGCTGCGGAGCGGAGCGGCGGCGGCGGCGGCA 443
DB |||||||
QY 582 GCGCTCGTCTGCTGCGCGGCGAGGCTGTGGAGCAGGAGCGGCGGAGCTGCGCGCACGC 641
DB |||||||
QY 444 GCGCTCGTCTGCTGCGCGGCGAGGCTGTGGAGCAGGAGCGGCGGAGCTGCGCGCACGC 503
DB |||||||
QY 642 CTTGCTGGTCTCTGCATCGAGACAGCGTCATGACCTCTTCTTCCAG 690
DB |||||||
QY 504 CTTGCTGGTCTCTGCATCGAGAACAGCGTCATGACCTCTTCTTCCAG 552
DB |||||||

RESULT 4

AA146065/C
ID AAL46065 standard; DNA; 35 BP.
XX
AC AAL46065;
XX
AT 19-JUL-2002 (first entry)
XX
DE Murine collagen XVIII PCR primer used to amplify canine collagen cDNA #1.
XX
KW Dog; pro-endostatin; endostatin; angiogenesis; cancer; gene therapy;
KW psoriasis; rheumatoid arthritis; retinopathy; macular degeneration;
KW corneal graft rejection; neovascular glaucoma; retrolental fibroplasia;
KW rubeosis; Osler-Webber Syndrome; myocardial angiogenesis;
KW plaque neovascularisation; telangiectasia; haemophilic joints;
KW angiofibroma; wound granulation; coronary collateral;
KW cerebral lateral; arteriovenous malformation;
KW ischaemic limb angiogenesis; diabetic neovascularisation; fracture;
KW cytostatic; antiarthritic; antipsoriatic; antidiabetic; ophthalmological;
KW gynaecological; mouse; collagen XVIII; PCR; primer; ss.
XX
OS Mus sp.
XX
PN EPI191036-A2.
XX
PD 27-MAR-2002.
XX
PF 24-AUG-2001; 2001EP-0307224.
XX
PR 25-AUG-2000; 2000US-227924P.
XX
PA (PFIZ) PFIZER PROD INC.
XX
PI Sheppard MG, Tong X;
XX
DR WPI; 2002-354068/39.
XX
PT An isolated nucleic acid molecule for the treatment of
PT angiogenesis-related disorder, such as cancers or diabetic retinopathy,
PT encodes an endostatin protein -
XX
PS Examples; Page 36; 56pp; English.
XX
CC The present invention provides the protein and coding sequences of canine
CC pro-endostatin and endostatin. The sequences can be used in the treatment
CC and diagnosis of angiogenesis related disorders, including cancer,
CC rheumatoid arthritis, psoriasis, retinopathy, macular degeneration,

CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasia,
CC rubeosis, Osler-Webber Syndrome, myocardial angiogenesis, plaque
CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,
CC wound granulation, coronary collaterals, cerebral collaterals,
CC arteriovenous malformations, ischaemic limb angiogenesis, diabetic
CC neovascularisation, and fractures. The present sequence is a PCR primer
CC derived from murine collagen XVIII cDNA and used to amplify the canine
CC collagen XVIII cDNA.
XX

SQ Sequence 35 BP; 5 A; 7 C; 5 G; 18 T; 0 other;

Query Match 4.2%; Score 35; DB 24; Length 35;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCTCTCAAGAAATAAAAGGAAGCCAAAGAG 829
DB |||||||
DB 35 ATGTAATCTCTCAAGAAATAAAAGGAAGCCAAAGAG 1

RESULT 5

ABN94215/C

ID ABN94215 standard; DNA; 251 BP.

XX AC

XX ABN94215;

XX DT

XX 13-AUG-2002 (first entry)

XX DE

XX Gene #713 used to diagnose liver cancer.

XX KW

XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX

OS Homo sapiens.

XX WO200229103-A2.

XX PD

XX 11-APR-2002.

XX PF

XX 02-OCT-2001; 2001WO-US30589.

XX PR

XX 02-OCT-2000; 2000US-237054P.

XX PA

XX (GENE-) GENE LOGIC INC.

XX PI

XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX WPI; 2002-426119/45.

XX DR

XX Claim 1; SEQ ID NO 713; 298pp; English.

XX CC

XX The invention relates to a novel method for diagnosing and detecting the

XX progression of liver cancer, hepatocellular carcinoma or metastatic liver

XX tumour in a patient, and differentiating metastatic liver cancer from

XX hepatocellular carcinoma in a patient, involving detecting the level of

XX expression of two or more genes represented in ABN93503-ABN97455 in a

XX tissue sample. The method of the invention has hepatotropic, and

XX cytostatic activity. The method is useful for diagnosing and detecting

XX the progression of liver cancer, hepatocellular carcinoma and metastatic

XX liver carcinoma in a patient. The method is useful for identifying

XX expression profiles which serve as useful diagnostic markers as well as

XX markers that can be used to monitor disease states, disease progression,

XX drug toxicity, drug efficacy and drug metabolism.

XX Note: The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX

```
SQ Sequence 251 BP; 44 A; 72 C; 75 G; 60 T; 0 other;
Query Match 4.2%; Score 35; DB 24; Length 251;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCCTCAAGAAATAAAGGAGCCCAAGAG 829
Dl |||||
41 ATGTAATCCTCAAGAAATAAAGGAGCCCAAGAG 7

RESULT 6
AAK53858/c
ID AAK53858 standard; cDNA; 338 BP.
XX
AC AAK53858;
XX
DT 16-NOV-2001 (first entry)
XX
DE Murine transport and binding associated protein encoding cDNA SEQ ID 423.
XX
KW Murine; liver; gene library; amino acid synthesis; binding protein;
KW cell metabolism; energy metabolism; fatty acid metabolism; synthesis;
KW phospholipid metabolism; purine; pyrimidine; nucleoside; nucleotide;
KW replication; transcription; translation; transport protein; ss.
XX
OS Mus musculus.
XX
PN DE20103510-UL.
XX
PD 07-JUN-2001.
XX
PF 28-FEB-2001; 2001DB-2003510.
XX
PR 02-DEC-1999; 99DB-1058160.
XX
PA (LION-) LION BIOSCIENCE AG.
XX
WPI; 2001-368570/39.
XX
Gene library containing sequences with specific 3'-ends and no polyA
tail, encoding proteins involved in a wide range of cellular processes
-
Claim 15; Page 134; 251pp; German.
XX
This invention describes a novel gene library (A) comprises a gene
sequence (or its part) encoding a protein involved in amino acid
synthesis, cellular/energy metabolism, metabolism of
fatty acids/phospholipids, synthesis or breakdown of
purines/pyrimidines/nucleosides/nucleotides, DNA
replication/transcription/translation, or is a transport/binding protein.
(A) are produced that correspond to the 3'-end of mRNA but without the
polyA tail. They can be prepared more efficiently and with less effort
than conventional libraries. AAK53436-AAK54275 represent fragments of the
gene library described in the method of the invention.
XX
SQ Sequence 338 BP; 71 A; 92 C; 83 G; 92 T; 0 other;
Query Match 4.2%; Score 35; DB 22; Length 338;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCCTCAAGAAATAAAGGAGCCCAAGAG 829
Dl |||||
39 ATGTAATCCTCAAGAAATAAAGGAGCCCAAGAG 5

RESULT 7
ABQ55585
ID ABQ55585 standard; cDNA; 697 BP.
XX
AC ABQ55585;
XX
```

22-AUG-2002 (first entry)
Human ovarian antigen HODFU24 cDNA, SEQ ID NO:1465.
Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
ovarian cancer; breast cancer; tumour; reproductive system disorder;
infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
inflammatory condition; immune disorder; blood disorder;
cardiovascular disorder; respiratory disorder; neurological disorder;
gastrointestinal disorder; urinary system disorder; drug screening;
gene therapy; chromosome mapping; forensic analysis;
antibody preparation; cytostatic; immunomodulatory; neuroprotective;
antiinflammatory; gynaecological; reproductive; gene; ss.
Homo sapiens.
WO200200677-A1.
03-JAN-2002.
07-JUN-2001; 2001WO-US18569.
07-JUN-2000; 2000US-209467P.
(HUMA-) HUMAN GENOME SCI INC.
Birse CE, Rosen CA;
WPI; 2002-147878/19.
P-PSDB; ABP42508.
Isolated nucleic acid molecules encoding novel ovarian polypeptides,
useful in the prevention, treatment and diagnosis of cancer (e.g.
ovarian cancer), immune disorders, cardiovascular disorders and
neurological diseases -
Claim 1; SEQ ID No 1465; 2922pp; English.
The invention relates to 2175 novel human ovarian antigens (ABP41054-
ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
encompasses polypeptides 90% identical and polynucleotides 95% identical
to the sequences of the invention. The invention additionally relates to
recombinant vectors and host cells comprising human ovarian antigen
polynucleotides, antibodies against human ovarian antigens, and the use
of ovarian antigen polynucleotides and polypeptides in diagnosing,
treating, prognosing or preventing various ovary and/or breast-related
disorders. Such conditions include ovarian cancer and breast cancer, and
metastatic tumours of ovarian or breast origin, reproductive system
disorders (e.g., infertility, disorders of pregnancy, anovulation,
polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
vaginitis), immune disorders (e.g., congenital and acquired
immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
blood-related disorders (e.g., anaemia), cardiovascular disorders,
respiratory disorders, neurological disorders, gastrointestinal disorders
and urinary system disorders. Ovarian antigen polypeptides and
polynucleotides may also be used in screening for compounds which
modulate ovarian antigen expression or activity. The polynucleotides may
further be used for gene therapy, chromosome mapping, in the
identification of individuals and in forensic analysis, and the
polypeptides may be used as food additives or to prepare antibodies
useful in disease diagnosis, drug targeting and phenotyping. The present
sequence represents cDNA encoding a human ovarian antigen of the
invention.
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.
Sequence 697 BP; 173 A; 205 C; 184 G; 114 T; 21 other;

```
Query Match      4.2%; Score 35; DB 24; Length 697;
Best Local Similarity 100.0%; Pred. No. 3.4e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      795 ATGTAATCCTCAAGAAATAAAAGGAGCCAAAGAG 829
      282 ATGTAATCCTCAAGAAATAAAAGGAGCCAAAGAG 316
Db

RESULT 8
AAT84484
ID AAT84484 standard; cDNA; 3394 BP.
XX
AC AAT84484;
XX
DT 19-NOV-1997 (first entry)
XX
DE Human alpha-1 collagen cDNA.
XX
KW Alpha-1 collagen; type XVIII collagen; cartilage degeneration; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2055
      /*tag= a
      /product= "alpha(XVIII) collagen"
      /note= "Partial sequence, no start codon given"
XX
PN US5643783-A.
XX
XX 01-JUL-1997.
XX
XX 01-DEC-1993; 93US-0159784.
XX
XX 01-DEC-1993; 93US-0159784.
XX
XX (HARD ) HARVARD COLLEGE.
XX
XX Oh SP, Olsen BR;
XX
XX WPI; 1997-350247/32.
XX
XX P-PSDB; AAW26327.
XX
XX Nucleic acid encoding human alpha-1 collagen - for production of
XX recombinant alpha-1 collagen, for use in the treatment of cartilage
XX degeneration
XX
XX Claim 1; Column 23-30; 35pp; English.
XX
XX A cDNA clone (AAT84484) codes for a human novel type alpha-1 (XVIII)
XX collagen (AAW26327) that is expressed in multiple tissues, especially
XX liver, lung and kidney. It was isolated from a placental cDNA
XX library using a probe based on an unidentified collagenous protein
XX and a probe based on mouse alpha-1 cDNA clone mc19. A claimed
XX plasmid comprising alpha-1 collagen nucleic acid and an expression
XX control sequence can be used to express recombinant collagen in
XX prokaryotic or eukaryotic (especially mammalian) host cells. The
XX collagen may be used to treat a patient suffering from a disease
XX associated with degradation of cartilage, and for supplementing
XX collagen.
XX
SQ Sequence 3394 BP; 668 A; 1130 C; 1080 G; 516 T; 0 other;

Query Match      4.2%; Score 35; DB 18; Length 3394;
Best Local Similarity 100.0%; Pred. No. 3.4e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      795 ATGTAATCCTCAAGAAATAAAAGGAGCCAAAGAG 829
      2187 ATGTAATCCTCAAGAAATAAAAGGAGCCAAAGAG 2221
Db

RESULT 9
AAX78379
```

```
ID AAX78379 standard; cDNA; 3394 BP.
XX
XX AAX78379;
XX
DT 25-AUG-1999 (first entry)
XX
DE Human alpha1 (XVIII) collagen cDNA.
XX
KW Alpha(XVIII) collagen; mimetic; endostatin; atomic coordinate; library;
KW anti-angiogenic; heparin binding domain; receptor binding domain; mimic;
KW alpha-helix A domain; carbohydrate recognition domain; CRD domain;
KW treatment; angiogenesis; tumour; human; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2055
      /*tag= a
      /product= "alpha(XVIII) collagen"
      /note= "Partial sequence, no start codon given"
XX
PN WO9931616-A1.
XX
XX 24-JUN-1999.
XX
XX 16-DEC-1998; 98WO-US26783.
XX
XX 16-DEC-1997; 97US-0069727.
XX
XX (HARD ) HARVARD COLLEGE.
XX
XX Hohenester E, Olsen BR, Sasaki T, Timpl R;
XX
XX WPI; 1999-395243/33.
XX
XX P-PSDB; AAY25113.
XX
XX Identifying mimetics of mammalian endostatin
XX
XX Disclosure; Fig 5A-C; 75pp; English.
XX
XX This invention describes a novel method for identifying mimetics of
XX mammalian endostatin. The method comprises identifying a compound
XX having atomic coordinates with non-trivial similarity to selected
XX coordinates of atoms of a mammalian endostatin involves (a) providing
XX a library of atomic coordinates of compounds in a library of candidate
XX compounds, (b) comparing the library of atomic coordinates to the
XX selected coordinates of a mammalian endostatin and (c) selecting from the
XX library at least one candidate compound on the basis of selection
XX criteria which include similarities between the atomic coordinates of the mammalian
XX endostatin. The invention also describes the use of an anti-angiogenic
XX fragment of endostatin comprising a domain selected from a heparin
XX binding domain, a receptor binding domain, and exposed on alpha-helix A
XX domain, and a carbohydrate recognition domain (CRD) domain. The methods
XX can be used for designing and selecting endostatin mimics. The compounds
XX identified can be used for treating undesired angiogenesis, e.g. tumours.
XX This sequence encodes human alpha1 (XVIII) collagen which is used in the
XX description of the method.
XX
SQ Sequence 3394 BP; 668 A; 1130 C; 1080 G; 516 T; 0 other;

Query Match      4.2%; Score 35; DB 20; Length 3394;
Best Local Similarity 100.0%; Pred. No. 3.4e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      795 ATGTAATCCTCAAGAAATAAAAGGAGCCAAAGAG 829
      2187 ATGTAATCCTCAAGAAATAAAAGGAGCCAAAGAG 2221
Db

RESULT 10
ABN95680
ID ABN95680 standard; DNA; 3394 BP.
```

XX AC ABN95680;
 XX DT 13-AUG-2002 (first entry)
 XX DE Gene #2178 used to diagnose liver cancer.
 XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 XX KW metastatic liver tumor; cytostatic; expression profile; disease state;
 XX KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 XX OS Homo sapiens.
 XX PN WO200229103-A2.
 XX PD 11-APR-2002.
 XX PF 02-OCT-2001; 2001WO-US30589.
 XX PR 02-OCT-2000; 2000US-237054P.
 XX PA (GENE-) GENE LOGIC INC.
 XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 XX DR WPI; 2002-426119/45.
 XX PT Diagnosing and detecting the progression of liver cancer,
 XX PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 XX PT involves detecting the level of expression of two or more genes in a
 XX PT liver tissue sample -
 XX PS Claim 1; SEQ ID NO 2178; 298pp; English.
 XX CC The invention relates to a novel method for diagnosing and detecting the
 XX CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 XX CC tumor in a patient, and differentiating metastatic liver cancer from
 XX CC hepatocellular carcinoma in a patient, involving detecting the level of
 XX CC expression of two or more genes represented in ABN93503-ABN97455 in a
 XX CC tissue sample. The method of the invention has hepatotropic, and
 XX CC cytostatic activity. The method is useful for diagnosing and detecting
 XX CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 XX CC liver carcinoma in a patient. The method is useful for identifying
 XX CC expression profiles which serve as useful diagnostic markers as well as
 XX CC markers that can be used to monitor disease states, disease progression,
 XX CC drug toxicity, drug efficacy and drug metabolism.
 XX CC Note: The sequence data for this patent did not form part of the printed
 XX CC specification, but was obtained in electronic format directly from WIPO
 XX CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 3394 BP; 668 A; 1130 C; 1080 G; 516 T; 0 Other;
 Query Match 4.2%; Score 35; DB 24; Length 3394;
 Best Local Similarity 100.0%; Pred. No. 3.4e-06;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 795 ATGTAATCTCTCAAGAAATAAAGGAGCCCAAGAG 829
 DB 2187 ATGTAATCTCTCAAGAAATAAAGGAGCCCAAGAG 2221
 RESULT 11
 AAT84485
 ID. AAT84485 standard; cDNA; 4031 BP.
 XX AC AAT84485;
 XX DT 19-NOV-1997 (first entry)
 XX DE Mouse alpha-1 collagen cDNA.
 XX KW Alpha-1 collagen; type XVIII collagen; cartilage degeneration; ss.
 XX OS Homo sapiens.

OS Mus musculus.
 XX FH Key Location/Qualifiers
 XX FT CDS 1..3867
 XX FT /*tag= a
 XX PN US5643783-A.
 XX PD 01-JUL-1997.
 XX PF 01-DEC-1993; 93US-0159784.
 XX PR 01-DEC-1993; 93US-0159784.
 XX PA (HARD) HARVARD COLLEGE.
 XX PI Oh SP, Olsen BR;
 XX DR WPI; 1997-350247/32.
 XX DR P-PSDB; AAW26328.
 XX PT Nucleic acid encoding human alpha-1 collagen - for production of
 XX PT recombinant alpha-1 collagen, for use in the treatment of cartilage
 XX PT degeneration
 XX PS Disclosure; Fig 2; 35pp; English.
 XX CC A cDNA clone (AAT84485) codes for a mouse novel type alpha-1 (XVIII)
 XX CC collagen (AAW26328) that is expressed in multiple organs, especially
 XX CC the liver, lung and kidney. It was isolated by screening murine
 XX CC 15.5- and 17.5-day embryo libraries with probes based on murine
 XX CC type XII collagen and on human alpha-1 collagen cDNA. Isolated
 XX CC nucleic acids can be used to express recombinant alpha-1 collagen
 XX CC in transformed host cells. Claimed nucleic acid (see AAT84484)
 XX CC coding for human alpha-1 collagen (AAW26327) can be used to treat
 XX CC patients suffering from diseases associated with degradation of
 XX CC cartilage, and for supplementing collagen.
 XX SQ Sequence 4031 BP; 873 A; 1172 C; 1266 G; 720 T; 0 other;
 Query Match 4.2%; Score 35; DB 18; Length 4031;
 Best Local Similarity 100.0%; Pred. No. 3.4e-06;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 795 ATGTAATCTCTCAAGAAATAAAGGAGCCCAAGAG 829
 DB 3987 ATGTAATCTCTCAAGAAATAAAGGAGCCCAAGAG 4021
 RESULT 12
 ABQ54955
 ID. ABQ54955 standard; cDNA; 4875 BP.
 XX AC ABQ54955;
 XX DT 22-AUG-2002 (first entry)
 XX DE Human ovarian antigen HEBK29 cDNA, SEQ ID NO:835.
 XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 XX KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 XX KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 XX KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 XX KW inflammatory condition; immune disorder; blood disorder;
 XX KW cardiovascular disorder; respiratory disorder; neurological disorder;
 XX KW gastrointestinal disorder; urinary system disorder; drug screening;
 XX KW gene therapy; chromosome mapping; forensic analysis;
 XX KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 XX KW antiinflammatory; gynaecological; reproductive; chromosome 21q22.3;
 XX KW gene; ss.
 XX OS Homo sapiens.

PN WO200200677-A1.
XX 03-JAN-2002.
XX 07-JUN-2001; 2001WO-US18569.
XX 07-JUN-2000; 2000US-209467P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Birse CE, Rosen CA;
XX WPI; 2002-147878/19.
XX P-PSDB; ABP41878.
XX isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX useful in the prevention, treatment and diagnosis of cancer (e.g.
XX ovarian cancer), immune disorders, cardiovascular disorders and
XX neurological diseases -
XX Claim 1; SEQ ID No 835; 2922pp; English.
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to
XX recombinant vectors and host cells comprising human ovarian antigen
XX polynucleotides, antibodies against human ovarian antigens, and the use
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX treating, prognosing or preventing various ovary and/or breast-related
XX disorders. Such conditions include ovarian cancer and breast cancer, and
XX metastatic tumours of ovarian or breast origin, reproductive system,
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX vaginitis), immune disorders (e.g., congenital and acquired
XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX respiratory disorders, neurological disorders, gastrointestinal disorders
XX and urinary system disorders. Ovarian antigen polypeptides and
XX polynucleotides may also be used in screening for compounds which
XX modulate ovarian antigen expression or activity. The polynucleotides may
XX further be used for gene therapy, chromosome mapping, in the
XX identification of individuals and in forensic analysis, and the
XX polypeptides may be used as food additives or to prepare antibodies
XX useful in disease diagnosis, drug targeting and phenotyping. The present
XX sequence represents cDNA encoding a human ovarian antigen of the
XX invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4875 BP; 892 A; 1569 C; 1570 G; 829 T; 15 other;

Query Match 4.2%; Score 35; DB 24; Length 4875;
Best Local Similarity 100.0%; Pred. No. 3.4e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCCTCAAGAAATAAAGGAAGCAAGAG 829
|||||
Db 3195 ATGTAATCCTCAAGAAATAAAGGAAGCAAGAG 3229
|||||

RESULT 13
AAX77720
ID AAX77720 standard; DNA; 5408 BP.
XX AAX77720;
XX
XX 10-AUG-1999 (first entry)
XX
XX Human collagen 18 DNA.

XX Plasminogen; human; angiotensin; endostatin; gene therapy; collagen 18;
XX anti-angiogenic; attenuation; cytostatic; anti-diabetic; ophthalmology;
XX tumour growth; solid tumour; diabetic retinopathy; retina; vector; ss.
XX Homo sapiens.
XX WO9926480-A1.
XX 03-JUN-1999.
XX 20-NOV-1998; 98WO-US24950.
XX 20-NOV-1997; 97US-0975424.
XX (GENE-) GENETIX PHARM INC.
XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX Bachelot T, Leboulch P, Pawliuk RJ;
XX WPI; 1999-357696/30.
XX P-PSDB; AAY08694.
XX Anti-angiogenic gene therapy vectors
XX Disclosure; Page 75-76; 83pp; English.
XX This invention describes a novel viral gene therapy vector comprising a
XX nucleic acid molecule encoding an anti-angiogenic polypeptide chosen
XX from human or murine angiotensin, human or murine endostatin and
XX angiogenesis-inhibiting fusions and fragments, where the viral vector is
XX sufficiently attenuated for use in human gene therapy. The products of
XX the invention have anti-angiogenic, cytostatic, anti-diabetic and
XX ophthalmological activity. The vector is used in gene therapy for
XX inhibiting tumour growth in humans harbouring a solid tumour. The vector
XX expresses an anti-angiogenic polypeptide. An additional use comprises
XX treatment of diabetic retinopathy, where the anti-angiogenic polypeptide
XX inhibits angiogenesis in the vicinity of the retina. The vector is
XX administered to cells ex vivo and then administered to the patient.
XX
SQ Sequence 5408 BP; 986 A; 1820 C; 1796 G; 806 T; 0 other;

Query Match 4.2%; Score 35; DB 20; Length 5408;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCCTCAAGAAATAAAGGAAGCAAGAG 829
|||||
Db 4202 ATGTAATCCTCAAGAAATAAAGGAAGCAAGAG 4236
|||||

RESULT 14
AAK83781/C
ID AAK83781 standard; DNA; 37664 BP.
XX AAK83781;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38593.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX WO200157182-A2.
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 02-OCT-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-483426/52.
 XX
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX

PS Disclosure; SEQ ID NO 38593; 3071bp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK51921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 37664 BP; 6667 A; 11245 C; 10958 G; 8794 T; 0 other;

Query Match 4.2%; Score 35; DB 22; Length 37664;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCCTCAAGAAATAAAGGAGCCCAAGAG 829
|||||
Db 29615 ATGTAATCCTCAAGAAATAAAGGAGCCCAAGAG 29581

RESULT 15
AAK83782/c
ID AAK83782 standard; DNA; 61710 BP.
XX
AC AAK83782;
XX
DT 07-NOV-2001 (first entry)
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38594.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232197.
PR 14-SEP-2000; 2000US-0232198.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -

XX Disclosure; SEQ ID NO 38594; 307lpp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins, and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX

SQ Sequence 61710 BP; 11560 A; 18542 C; 18148 G; 13460 T; 0 other;
Query Match 4.2%; Score 35; DB 22; Length 61710;
Best Local Similarity 100.0%; Pred. No. 2.9e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 ATGTAATCCTCAAGAAATAAAAGGAAGCCAAAGAG 829
|||||
Db 29908 ATGTAATCCTCAAGAAATAAAAGGAAGCCAAAGAG 29874

Search completed: August 17, 2003, 16:43:15
Job time : 321.86 secs

THIS PAGE BLANK (USPTO)

; CURRENT APPLICATION NUMBER: US/09/449,293
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Rattus rattus
US-09-449-293-3

Query Match 45.4%; Score 376.2; DB 3; Length 558;
Best Local Similarity 79.7%; Pred. No. 1.9e-62;
Matches 444; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
QY 137 CCACACCCACAGGACTTCCAGCTGGTGTGACCTGGTGGCCCTGGAACAGCCCGCAGC 196
DB 2 CTCATCTCATAGGACTTTCAGCCAGTGTCTCACCTGGTGGACTGAACACCCCTGT 61
QY 197 CGGGCGGCATGCGAGGCATCCGGGAGCGGACTTCCAGTGTCTTCCAGCAGGCGCGCGCG 256
DB 62 CTGAGGCATCGTGTATCCGTGGAGCAGATTTCCAGTGTCTTCCAGCAAGCCCGAGCG 121
QY 257 CGGGGTGGCGGCACCTTCGGGGCTTCTGTCTGCGGGCTGCGAGGACTCTACAGCA 316
DB 122 TGGGGCTGTGGGCACCTTCGGGGCTTCTGTCTCTAGGCTGCGAGGATCTCTATAGCA 181
QY 317 TCGTGGCGCGCGCGGACCGGCGGTGCGGTCTGCTCAACCTCAGGGACGAGGTCTCT 376
DB 182 TCGTGGCGCGGTGTGACGGGGCTGTGTGCCATCTGTAACCTGAAGGACGAGGTCTAT 241
QY 377 TCCAGCTGGGAGGCTTATTTCTCGGGCTCCGAGGCCAGCTGGAAGCCGGGGCCCGCA 436
DB 242 CTCAGCTGGGACTCCCTGTTTCTGGCTCCAGGGTCAACTGCAACCCGGGGCCCGCA 301
QY 437 TCTTCTTTTCAGCGGAGAGATGTCTGACGACCCCGCTGGCCCGCCCGAAGAGGCTGT 496
DB 302 TCTTTTCTTTTGACGGCAGAGATGTCTGAGACACCCAGCTGGCGCGCAGAGGCTAT 361
QY 497 GGCAGCGCTCCGACCCCGGCGCGCTGACCGACAGCTACTCGAGAGAGTGGCGGA 556
DB 362 GGCAGCGCTCGACCCCGGCGGAGGCTGTATGGAGAGTACTGTGAGACATGGCGAA 421
QY 557 CGAGGCGCGCGCGCACCGGCGAGCGTGTCTGCTGTGGGGCGGAGCTGCTGGAGC 616
DB 422 CTGAACACTTGGGGCTACAGTCAAGCTCCTCCCTGTCTGAGCAGGCTCTTGAAC 481
QY 617 AGAGGCGCGGAGCTGCGGCGACGCTTCTGTGTCTGTGATCGAGAACAGGCTCATGA 676
DB 482 AGAAGCTGCGAGCTGCCAACACAGCTACATCTGCTGTGATGAGATAGCTTCATGA 541
QY 677 CCTCTTCTCCAAAGTAG 693
DB 542 CCTCTTCTCCAAATAG 558

RESULT 12
US-09-775-325-3
; Sequence 3, Application US/09775325
; Patent No. 6500449
; GENERAL INFORMATION:
; APPLICANT: Abitbol, Marc
; APPLICANT: Uteza, Yves
; APPLICANT: Menasche, Maurice
; APPLICANT: Bossard, Carine
; APPLICANT: Van Den Bergh, Loic
; APPLICANT: Bonnel, Sebastian
; APPLICANT: Prats, Hervé
; APPLICANT: Honiger, Jiri
; APPLICANT: Neuner-Jehle, Martin
; TITLE OF INVENTION: INTRAOCULAR TRANSPLANTATION OF ENCAPSULATED CELLS
; FILE REFERENCE: 8076.202USD1
; CURRENT APPLICATION NUMBER: US/09/775,325
; CURRENT FILING DATE: 2001-02-01

; PRIOR APPLICATION NUMBER: US/09/449,293
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Rattus rattus
US-09-775-325-3

Query Match 45.4%; Score 376.2; DB 4; Length 558;
Best Local Similarity 79.7%; Pred. No. 1.9e-62;
Matches 444; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
QY 137 CCACACCCACAGGACTTCCAGCTGGTGTGACCTGGTGGCCCTGGAACAGCCCGCAGC 196
DB 2 CTCATCTCATAGGACTTTCAGCCAGTGTCTCACCTGGTGGACTGAACACCCCTGT 61
QY 197 CGGGCGGCATGCGAGGCATCCGGGAGCGGACTTCCAGTGTCTTCCAGCAGGCGCGCGCG 256
DB 62 CTGAGGCATCGTGTATCCGTGGAGCAGATTTCCAGTGTCTTCCAGCAAGCCCGAGCG 121
QY 257 CGGGGTGGCGGCACCTTCGGGGCTTCTGTCTGCGGGCTGCGAGGACTCTACAGCA 316
DB 122 TGGGGCTGTGGGCACCTTCGGGGCTTCTGTCTCTAGGCTGCGAGGATCTCTATAGCA 181
QY 317 TCGTGGCGCGCGCGGACCGGCGGTGCGGTCTGCTCAACCTCAGGGACGAGGTCTCT 376
DB 182 TCGTGGCGCGGTGTGACGGGGCTGTGTGCCATCTGTAACCTGAAGGACGAGGTCTAT 241
QY 377 TCCAGCTGGGAGGCTTATTTCTCGGGCTCCGAGGCCAGCTGGAAGCCGGGGCCCGCA 436
DB 242 CTCAGCTGGGACTCCCTGTTTCTGGCTCCAGGGTCAACTGCAACCCGGGGCCCGCA 301
QY 437 TCTTCTTTTCAGCGGAGAGATGTCTGACGACCCCGCTGGCCCGCCCGAAGAGGCTGT 496
DB 302 TCTTTTCTTTTGACGGCAGAGATGTCTGAGACACCCAGCTGGCGCGCAGAGGCTAT 361
QY 497 GGCAGCGCTCCGACCCCGGCGCGCTGACCGACAGCTACTCGAGAGAGTGGCGGA 556
DB 362 GGCAGCGCTCGACCCCGGCGGAGGCTGTATGGAGAGTACTGTGAGACATGGCGAA 421
QY 557 CGAGGCGCGCGCGCACCGGCGAGCGTGTCTGCTGTGGGGCGGAGCTGCTGGAGC 616
DB 422 CTGAACACTTGGGGCTACAGTCAAGCTCCTCCCTGTCTGAGCAGGCTCTTGAAC 481
QY 617 AGAGGCGCGGAGCTGCGGCGACGCTTCTGTGTCTGTGATCGAGAACAGGCTCATGA 676
DB 482 AGAAGCTGCGAGCTGCCAACACAGCTACATCTGCTGTGATGAGATAGCTTCATGA 541
QY 677 CCTCTTCTCCAAAGTAG 693
DB 542 CCTCTTCTCCAAATAG 558

RESULT 13
US-09-252-991A-6774
; Sequence 6774, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6774
; LENGTH: 741

; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6774

Query Match 7.9%; Score 65.6; DB 4; Length 741;
Best Local Similarity 43.9%; Pred. No. 0.00027;
Matches 281; Conservative 0; Mismatches 359; Indels 0; Gaps 0;

QY 18 CATCTGCGCGCGCCCGCGCTGCTGGACCCCGCAGCCCTACCCCGGGCGCCCGCACCA 77
DB 64 CATCCCGCGCCCTTCCGCGCGAGCGCTGACCAACGTCTCTCGGCGAGTACCT 123
QY 78 CGGTCTCTAGTCACATTCAGCGCGCTCGCCCGCTGCTGGCGCTGCTCCACACCCACAC 137
DB 124 GGCACAGCATGGCAGACCCAGCTGCGCATGCGCGAGACCGAGAGTACGCCCATGTGAC 183
QY 138 CCACACCCACAGGACTTCAGCTGTGTGCTGCACTGTGTGCGCCCTGAACACCCCGCAGCC 197
DB 184 CTTCTTCTTCTCGCGCGCGCGAGGAACTTACGAGGCGGAGAGCGCATCTGTATCCC 243
QY 198 GGGGGGATGGAGGATTCGGGGAGCGGATTCAGTTCAGTTCAGGAGCGCGCGCGC 257
DB 244 CTCGCCAAGGTGCGCCACCTACGACCTGCAACCCGAGATGAGCGCGCGGAACTCACCGA 303
QY 258 GGGGCTGGCGGACCTTCGGGCGCTTCTGTCGTCGCGCTGCGAGACCTCTACAGCAT 317
DB 304 CCGTATCGTGAAGCCATCAGCAGCAGCGCTAGGAGTGTGTGTAACCTACGCCAA 363
QY 318 CGTGGCGCGCGCGACCGCGGTGCGCGTGTCTCAACTCAGGAGCAGGCTGTCTTT 377
DB 364 CGCGCATGCTGGCGCGCACCGCGGTGTTGAGGCGCGGTCAAGCGCTGGAGTGCCT 423
QY 378 CCCAGCTGGAGGCTTATTCGGGCTCGAGGCGCAGTGAAGCCCGGGCGCGCAT 437
DB 424 GGAACCTGATGGCGCGTATCTGTCGAGGCGCTGGACAAGGTGCGGCGCGAAGCCCTGAT 483
QY 438 CTTCTCTTTTCAGCGGAGAGATGCTGTCAGACCGCGCTGCGCGCGGCGGAGCGGTG 497
DB 484 CACGCGCGACACCGGCAATGTGAGCAGATGGAGACGATCCACCGCGCGCGGCGCAC 543
QY 498 GCACGGCTCCGACCCCGCGCGCGCGCTGACCGCAGCTACTGCGAGACGTCGGCGGAC 557
DB 544 CGCGCACCTGCGAACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 603
QY 558 GGAGGCG 617
DB 604 CGAAGCGCGGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 663
QY 618 GGAGGCGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657
DB 664 GCCGCGGAAATGACCGGTGCGAGCATCTGTCACCCCTCGGC 703

RESULT 14
US-09-252-991A-6920/c
; Sequence 6920, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6920
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-6920

Query Match 7.9%; Score 65.6; DB 4; Length 1557;
Best Local Similarity 43.9%; Pred. No. 0.00027;
Matches 281; Conservative 0; Mismatches 359; Indels 0; Gaps 0;

QY 18 CATCTGCGCGCGCCCGCGCTGCTGGACCCCGCAGCCCTACCCCGGGCGCCCGCACCA 77
DB 793 CATCCCGCGCCCTTCCGCGCGAGCGCTGACCAACGTCTCTCGGCGAGTACCT 734
QY 78 CGGTCTCTAGTCACATTCAGCGCGCTCGCCCGCTGCTGGCGCTGCTCCACACCCACAC 137
DB 733 GGCACAGCATGGCAGACCCAGCTGCGCATGCGCGAGACCGAGAGTACGCCCATGTGAC 674
QY 138 CCACACCCACAGGACTTCAGTGTGTGCTGACCTGTGTGCGCCCTGAACAGCCCGCAGCC 197
DB 673 CTTCTTCTTCTCGCGCGCGCGAGGAACTTACGAGGCGCGAGGAGCGCATCTGTATCCC 614
QY 198 GGGGCGATGCGAGGATCGCGGCGCGCTTCCAGTGTCTTCCAGCAGCGCGCGCGCGC 257
DB 613 CTCGCCAAGGTGCGCCACCTACGACCTGCAACCCGAGATGAGGCGCGCGGAGTACCGA 554
QY 258 GGGGCTGGCGCGCACCTTCGGGCGCTTCTGTCGTCGCGCTGCGAGACCTCTTACAGCAT 317
DB 553 CCGTATCGTGAAGCCATCAGCAGCAGCGCTAGCAGCTGATCGTGTCAACTACGCCAA 494
QY 318 CGTGGCGCGCGCGACCGCGGTGCGCGTGTCTCAACTCAGGAGCAGGCTGTCTTT 377
DB 493 CGCGCATGCTGGCGCGCACCGCGGTGTTGAGGCGCGCGGTCAAGCGCTGGAGTGCCT 434
QY 378 CCCAGCTGGAGGCGCTTATTCGGGCTCCGAGGCGCAGCTGAAGCCCGGGCGCGCAT 437
DB 433 GGAACCTGATGGCGCGTATCTGTCAGGCGCTGGACAAGTTCGGCGCGGAGCCCTGAT 374
QY 438 CTTCTCTTTTCAGCGGAGAGATGCTGTCAGCAGCAGCGCTGCGCGCGCGCGCGCGG 497
DB 373 CACCGCGCACCGCAATGTGAGCAGATGGAAGACGAGTCCACCGCGCGCGCGCACAC 314
QY 498 GCACGGCTCCGACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 557
DB 313 CGCGCACCTGCGAACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 254
QY 558 GGAGGCG 617
DB 253 CGAAGCGCGGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 194
QY 618 GGAGGCGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657
DB 193 GCCGCGGAAATGACCGGTGCGAGCATCTGTCACCCCTCGGC 154

RESULT 15
US-09-252-991A-6824
; Sequence 6824, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6824
; LENGTH: 1590
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6824

Query Match	7.9%; Score 65.6; DB 4; Length 1590;
Best Local Similarity	43.9%; Pred. No. 0.00027;
Matches 281; Conservative	0; Mismatches 359; Indels 0; Gaps 0;
Qy	18 CATCCTGGCGGGCCCCCGGCGCTGTCTGGACACCCCGACCCCTACCCCGGGGCCCGGACCA 77
Db	948 CATCCGGCGCCCTTGGCCCTTCCCGCGAGCGCTGACCAACGTCTCTCGCGAGTACCT 1007
Qy	78 CGGCTCTACGTGCACTTCCAGCCGGCTCGCCCACTGTGTGGGCCGTGCACACCCAC 137
Db	1008 GGGCAAGCATGGCAAGACCCAGCTGCGCATCGCCGAGACCGAGAAGTACGCCCATGTGAC 1067
Qy	138 CCACACCCACAGGACTTCCAGCTGTGTGTGCACCTGTGTGCCCTGAACAGCCCGGAGCC 197
Db	1068 CTTCCTTCTTCGGCGGCCCGAGGAACCTTACGAGGCGGAGAGCGCATCTGTATCCC 1127
Qy	198 GGGCGGCGATCGAGAGCATCCGGGAGCGGACTTCCAGTGTCTTCAGCAGCGCGCGCCGC 257
Db	1128 CTCGCCAAGGTGCGCACTTACGACCTGTCAACCGAGATGAGCGCGCGGAAGTCAACGA 1187
Qy	258 GGGGCTGGCGGCACCTTTCGGGCGCTTCTGTGCTCGTCCGGCTGCAGAGCCTCTACAGCAT 317
Db	1188 CCGTATCGTCGAGCCATCGAGCAACAGCGCTACGACGTGATCTGTGTCAACTACGGCAA 1247
Qy	318 CGTGCGCCGGCGCGACCGCACCGGGGTGCCGCTCGTCAACCTCAGGAGCAGAGTGTCTTT 377
Db	1248 CGGCGCATGTCGGCCACACCGGGGTGTTTCGAGGCCCGGTCAGAGCCGTGGAGTGCCCT 1307
Qy	378 CCCAGTCTGGAGGCGCTTATTTCTCGGCTCGAGGGCCAGCTGAGCCCGGGGCCCGCAT 437
Db	1308 GGCACCTGTATGGGCGCTATCGTCTGAGGGCGCTGGACAAGGTCTCGGCGCGAAGCCCTGAT 1367
Qy	438 CTTCCTTTTCGACGGCAGAGATGCTCTGCAGACCCCGCTGCGCCCGGAAGAGCGTGTG 497
Db	1368 CACGGCGCACCA CGGCAATGTCGAGCAGATGGAAGACAGTCCACCGCCAGGGCGACAC 1427
Qy	498 GCAAGGCTCCGACCCCAAGCGGGCGCCGCTGACCGACAGCTATCTGCAGACGTGGGGGAC 557
Db	1428 CGCGCACACCTGGGAAACCGGTGCGCTTCTAGCTCGGAAGCGCAAGCTGAGCATCCG 1487
Qy	558 GGAAGGCCCGGGCGGCACCGGGCAGGGCGTCTGTCTGCTGTCGGGGCAGGCTGCTGGAGCA 617
Db	1488 CGAAGGGCGGTGCTGSCCGACGTGGCGCCGACCATGCTGACCTCATGGGGCTGGAGCA 1547
Qy	618 GGAAGGCCCGAGTGGCGCCACGCCCTTCTGTGCTCTGCG 657
Db	1548 GCCGGCGGGAATGACCCGGTTCGACGATCGTCACTCCGCG 1587

THIS PAGE BLANK (USPTO)

2000

Db 1494 -----CCACCCCGCCACAGCCACCGCGACTTCCAGCCGGTGCTCCACCTGGTGGG 1545
QY 181 CTGAACAGCCCGCAGCCGGCGCATGCGAGGATCCGGGGAGCGGACTTCCAGTGTCTC 240
Db 1546 CTCACAGCCCGCTGTACGGCGCATGCGGGGATCCGGGGGCGGACTTCCAGTGTCTC 1605
QY 241 CAGCAGCGCGCGCGCGCGGGTGGCGGCACCTTCCGGGCGCTTCTGTGCTCGCGGCTG 300
Db 1606 CAGCAGCGCGCGCGCTGGGGTGGCGGCACCTTCCGGCGCTTCTGTGCTCGCGCTG 1665
QY 301 CAGGACCTCTACAGCATGTGTGCGCGCGCGCAGCCGACCGGGGTGCGCGCTGTCTGAACCTC 360
Db 1666 CAGGACCTGTACAGCATGTGTGCGCGCGCGCAGCCGCGCGCGCTGTCTGAACCTC 1725
QY 361 AGGACGAGGTGCTCTTCCCGAGCTGGGAGGCTTATCTCGGGCTCCGAGGGCGAGCTG 420
Db 1726 AAGGACGAGGTGCTGTGTTTCCAGCTGGGAGGCTCTGTCTCAAGGCTCTGAGGGTCCGCTG 1785
QY 421 AAGCCCGGGCGCGCATCTTCTTCTTCGACGGCAGAGATGTCTCGACACCCCGCTGG 480
Db 1786 AAGCCCGGGCGCGCATCTTCTTCTTCGACGCAAGGAGCTCTGAGGCACCCCGCTGG 1845
QY 481 CCCCAGAAAGCGGTGTGGCAAGGCTCCGACCCCGCGGGCGCGCTGTGACGACAGCTAC 540
Db 1846 CCCCAGAAAGCGGTGTGGCATGGCTCGGACCCCGCGCGCGAGGCTGACCCGAGAGCTAC 1905
QY 541 TCGGACGCTGGGAGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 1906 TGTGACGCTGGGAGCGAGGCTCCCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1965
QY 601 GGCAGGCTGTGAGCAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 1966 GGCAGGCTGTGGGAGGAGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2025
QY 661 GAGAACAGCGTATGACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720
Db 2026 GAGAACAGCTTATGATGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2085
QY 721 GAGGCGCGCGCGCGAGGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 756
Db 2086 ACCGCGCGCGCGCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2145
QY 757 --GGGCTGTGCGCGCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 813
Db 2146 CAGGACCTGGCTGCCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2205
QY 814 AAGGAAGCCAAAGAG 829
Db 2206 AAGGAAGCCAAAGAG 2221

RESULT 4
US-10-060-036-144
; Sequence 144, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144
; LENGTH: 4551
; TYPE: DNA

; ORGANISM: Homo sapiens
US-10-060-036-144
Query Match 59.5%; Score 493.2; DB 14; Length 4551;
Best Local Similarity 83.9%; Pred. No. 5.6e-107; Indels 12; Gaps 1;
Matches 574; Conservative 0; Mismatches 98;
QY 1 CCCTGGCGGCGAGATGACATCTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db 3874 CCCTGGCGGCGAGATGACATCTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3933
QY 61 CCGGGGGCGCGCAGCAGCGCTCTACGTGCACTTCCAGCGCGCTCGCGCGCGCGCGCGCG 120
Db 3934 CCGGAGCGCGCGCAGCAGCTCTACGTGCACTTCCAGCGCGCGCGCGCGCGCGCGCG 3989
QY 121 CCGGTCCACACCCACACCCACACAGGACTTCCAGCTGTGCTGTGCTGTGCTGTGCTG 180
Db 3990 -----CCCAACCG 4041
QY 181 CTGAACAGCG 240
Db 4042 CTCACAGCGCGCGCTGTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4101
QY 241 CAGCAGGCG 300
Db 4102 CAGCAGGCG 4161
QY 301 CAGGACCTCTACAGCATGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db 4162 CAGGACCTGTACAGCATGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4221
QY 361 AGGAGCAGGCTGTCTTCCCGAGCTGGGAGGCTTATCTTCGCGGCTCCGAGGGCGCG 420
Db 4222 AAGGAGGCTGTCTTCCCGAGCTGGGAGGCTTCTGTCTCAGGCTCTGAGGGTCCGCTG 4281
QY 421 AAGCGCGGGCGCGCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 480
Db 4282 AAGCGCGGGCGCGCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4341
QY 481 CCCCAGAGAGCGGTGTGGCAGCGCTCCGACCCCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 4342 CCCCAGAGAGCGGTGTGGCATGGCTCGGACCCCGCGCGCGCGCGCGCGCGCGCGCGCG 4401
QY 541 TCGAGACCTGTGCGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 4402 TGTGAGCTGTGCGGAGCGAGGCTCCCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 4461
QY 601 GGCAGGCTCTGAGCAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 4462 GGCAGGCTCTGGGCGAGAGTGCAGGAGTGCAGGAGTGCAGGAGTGCAGGAGTGCAGG 4521
QY 661 GAGAACAGCGTATGACCTCTTC 684
Db 4522 GAGAACAGCTTACGCTGCCTCC 4545

RESULT 5
US-10-042-347-4
; Sequence 4, Application US/10042347
; Publication No. US20030114370A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Endostatin Protein and Peptide Fr
; FILE REFERENCE: 05213-0880 (43170-249874)
; CURRENT FILING DATE: 2002-01-11
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 60/106,343
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 09/154,302

;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: US 08/740,168
;; PRIOR FILING DATE: 1996-10-22
;; PRIOR APPLICATION NUMBER: US 60/005,835
;; PRIOR FILING DATE: 1995-10-23
;; PRIOR APPLICATION NUMBER: US 60/023,070
;; PRIOR FILING DATE: 1996-08-02
;; PRIOR APPLICATION NUMBER: US 60/026,263
;; PRIOR FILING DATE: 1996-09-17
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 4
;; LENGTH: 546
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-042-347-4

Query Match 50.4%; Score 418; DB 14; Length 546;
Best Local Similarity 85.3%; Pred. No. 2.9e-89;
Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY	139	CACACCCACGAGACTTCAGCTGGTGTCTGCTGACCTGGTGGCCCTGAACAGCCGCGAGCGG	198
Db	1	CACAGCACCGGACACTTCAGCCGGTGTCTCACCTGGTGGCTCAACAGCCCTGTCA	60
QY	199	GGCGGATGCGAGGACATCCGGGAGCGGACTTCCAGTGTCTTCAGCAGGCGCGCGCGG	258
Db	61	GGCGGATGCGGGGACATCCGGGGGCGGACTTCCAGTGTCTTCAGCAGGCGGGCGGTG	120
QY	259	GGGCTGGCGGACCTTCCTGGGCGCTTCCTGTCTGCGGGCTGAGGACCTTACAGCATC	318
Db	121	GGGCTGGCGGACCTTCCTGGGCGCTTCCTGTCTGCGGGCTGAGGACCTTACAGCATC	180
QY	319	GTGCGGCGCGGACCGGACCGGGGTGCGCGTCTCAACTCAGGACGAGTGTCTTTC	378
Db	181	GTGCGGCGGACCGGACCGGCGGCTGCGGCTCAACTCAGGACGAGTGTCTTTC	240
QY	379	CCAGCTGGGAGGCTTATTTCTCGGGCTCCGAGGCGGAGCTGAAGCCCGGGGCCGATC	438
Db	241	CCAGCTGGGAGGCTTGTCTCAGGCTCTGAGGCTCCGCTGAAGCCCGGGGCCGATC	300
QY	439	TTCTCTTTTTCAGCGCAGAGATGTCTTCAGCAGCACCCTGCGCCCGGAGAGCGTGTG	498
Db	301	TTCTCTTTTTCAGCGCAGAGATGTCTTCAGCAGCACCCTGCGCCCGGAGAGCGTGTG	360
QY	499	CAGGCTCCGACCCCGAGCGGCGCGCTGACCGAGCTACTGCGAGACGTGGCGGACG	558
Db	361	CATGGCTCGGACCCCAACCGGGCGAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGACG	420
QY	559	GAGGCCCGGCGGCGGACCGGCGGAGGCTGCTGCTGCTGCGGGGAGGCTGTGGAGCAG	618
Db	421	GAGGCTCCCTCGGCGGACCGGCGGAGGCTTCTGCTGCTGGGGGAGGCTTCTGGGGCAG	480
QY	619	GAGGCCGAGCTGCGGCGGACCGCTTGTGGTGTCTGTGATCGAGAACAGCGTCATGACC	678
Db	481	AGTGCCGAGCTGCCATCAGCGCTACATCGTGTCTGTGATTCGAGAACAGCTTCATGACT	540
QY	679	TCCTTC 684	
Db	541	GCCTCC 546	

RESULT 6

US-10-292-418-3
; Sequence 3, Application US/10292418
; Publication No. US20030139365A1
; GENERAL INFORMATION:
; APPLICANT: Lo, Kin-Ming
; APPLICANT: Li, Yue
; APPLICANT: Gillies, Stephen D
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
; TITLE OF INVENTION: Immunofusins
; FILE REFERENCE: LEX-006C1

;; CURRENT APPLICATION NUMBER: US/10/292,418
;; CURRENT FILING DATE: 2002-11-12
;; PRIOR APPLICATION NUMBER: 09/383,315
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: US 60/097,883
;; PRIOR FILING DATE: 1998-08-25
;; NUMBER OF SEQ ID NOS: 54
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 3
;; LENGTH: 549
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1)..(549)
;; OTHER INFORMATION: endostatin
US-10-292-418-3

Query Match 50.4%; Score 418; DB 12; Length 549;
Best Local Similarity 85.3%; Pred. No. 2.9e-89;
Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY	139	CACACCCACGAGACTTCAGCTGGTGTGCTGACCTGGTGGCCCTGAACAGCCGCGAGCGG	198
Db	1	CACAGCACCGGACACTTCAGCCGGTGTCTCACCTGGTGGCTCAACAGCCCTGTCA	60
QY	199	GGCGGATGCGAGGACATCCGGGAGCGGACTTCCAGTGTCTTCAGCAGGCGCGCGCGG	258
Db	61	GGCGGATGCGGGGACATCCGGGGGCGGACTTCCAGTGTCTTCAGCAGGCGGGCGGTG	120
QY	259	GGGCTGGCGGACCTTCCTGGGCGCTTCCTGTCTGCGGGCTGAGGACCTTACAGCATC	318
Db	121	GGGCTGGCGGACCTTCCTGGGCGCTTCCTGTCTGCGGGCTGAGGACCTTACAGCATC	180
QY	319	GTGCGGCGGCGGACCGGACCGGGGTGCGCGTCTCAACTCAGGACGAGTGTCTTTC	378
Db	181	GTGCGGCGGTCGCGACCGGCGGAGCGGCTGCGCATCTCAACTCAAGGACGAGTGTCTT	240
QY	379	CCAGCTGGGAGGCTTATTTCTCGGGCTCCGAGGCGGAGCTGAAGCCCGGGGCCGATC	438
Db	241	CCAGCTGGGAGGCTTGTCTCAGGCTCTGAGGCTCCGCTGAAGCCCGGGGCCGATC	300
QY	439	TTCTCTTTTTCAGCGCAGAGATGTCTTCAGCAGCACCCTGCGCCCGGAGAGCGTGTG	498
Db	301	TTCTCTTTTTCAGCGCAGAGATGTCTTCAGCAGCACCCTGCGCCCGGAGAGCGTGTG	360
QY	499	CAGGCTCCGACCCCGAGCGGCGCGCTGACCGAGCTACTGCGAGACGTGGCGGACG	558
Db	361	CATGGCTCGGACCCCAACCGGGCGAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGACG	420
QY	559	GAGGCCCGGCGGCGGACCGGCGGAGGCTGCTGCTGCTGCGGGGAGGCTGTGGAGCAG	618
Db	421	GAGGCTCCCTCGGCGGACCGGCGGAGGCTTCTGCTGCTGGGGGAGGCTTCTGGGGCAG	480
QY	619	GAGGCCGAGCTGCGGCGGACCGCTTGTGGTGTCTGTGATCGAGAACAGCGTCATGACC	678
Db	481	AGTGCCGAGCTGCCATCAGCGCTACATCGTGTCTGTGATTCGAGAACAGCTTCATGACT	540
QY	679	TCCTTC 684	
Db	541	GCCTCC 546	

RESULT 7

US-10-131-241-53
; Sequence 53, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565

```

; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-241-53

```

Query Match	50.4%;	Score 418;	DB 14;	Length 549;
Best Local Similarity	85.3%;	Pred. No. 2.9e-89;		
Matches 466;	Conservative 0;	Mismatches 80;	Indels 0;	Gaps 0;
QY	139	CACACCCACAGGACTTCCAGCTCGTGCACCTGTTGGCCCTGAACAGCCCGCAGCCG	198	
Db	1	CACAGCCACCGGACTTCCAGCCGGTGTCTCACTCTGTTGCCTCAACAGCCCCCTGTCA	60	
QY	199	GGCGGCATCGAGGCATCCGCGGAGCGGACTTCCAGTCTCTTCCAGCAGGCGCGCCGCG	258	
Db	61	GGCGCATCGGGCATCCGCGGGCCGACTTCCAGTGTCTTCCAGCAGGCGCGGCCGTG	120	
QY	259	GGGCTGGCGGCACCTTTCGGGGCCCTTCTGTCTGTCTGGCGGTGCAAGACCTCTACAGATC	318	
Db	121	GGGCTGGCGGGCACCTTTCGGGGCCCTTCTGTCTCTGGCGCTTGCAGGACCTGTATCAGCATC	180	
QY	319	GTGCGCCGCGCGACCCGACACCGGGGTGCCGTCTCAACTCTCAGGACGAGGTCTCTTTC	378	
Db	181	GTGCGCGTGGCGACCGGGCAGCCGTGCCCATCTGTCACTCAAGGACGAGTCTGTGTT	240	
QY	379	CCAGCTGGGAGGCTTATTTCTCGGGTCTCGAGGGCCAGTGAAGCCCGGGGCCCGCATC	438	
Db	241	CCCAGCTGGGAGGCTCTGTCTCAGGCTCTGAGGCTCCGCTGAAGCCCGGGGCACGCATC	300	
QY	439	TTCTCTTTTCACGGCAGAGATGTCCTCAGCACCCCGCCTGGCCCCCGGAGAGAGCGTGTGG	498	
Db	301	TTCTCTTTTCACGGCAAGAGAGTCTTAGGGCACCCACCTGGCCCCCAAGAGAGCGTGTGG	360	
QY	499	CACGGCTCCGACCCACGCGGGCGCGCTGTACCGACAGCTACTCGCAGACGTGGCGGACG	558	
Db	361	CATGGCTCGGACCCCAACGGGCGCAGGCTGACCGAGAGCTACTGTGAGAGCTGGCGGACG	420	
QY	559	GAGGCCCGGGGGCCACCGGGCAGGCGTGTCTGCTGTGGCGGGCAGGCTCTGTGAGGAG	618	
Db	421	GAGGCTCTCTCGGGCACGGGCGCAGGCGCTCTGCTGTGGGGGGCAGGCTCTCTGGGGCAG	480	
QY	619	GAGGCCCGGAGCTGCCGCCACCGCTCTCGTGTGTCTGTGCATCTCGAGAACAGGCTCATGACC	678	
Db	481	AGTGCCCGGAGCTGCCATCAGCCCTACATCGTGTCTGTGCATTGAGAACAGCTTCATGACT	540	
QY	679	TCCTTC	684	
Db	541	GCCTCC	546	

```

; RESULT 8
; US-09-873-676-30
; Sequence 30, Application US/09873676
; Patent No. US20020077289A1
; GENERAL INFORMATION:
; APPLICANT: MacDonald, Nicholas J.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use
; FILE REFERENCE: 05213-0378 (43170-259333)
; CURRENT APPLICATION NUMBER: US/09/873,676
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065

```

```

; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 30
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-676-30

Query Match          50.4%; Score 418; DB 9; Length 552;
Best Local Similarity 85.3%; Pred. No. 2.9e-89;
Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0

QY 139 CACACCACACAGAGACTTCACAGCTGTGTCTGCACCTGTGTGGCCCTGAAACAGCCCGCAGCGC 198
DB 1 CACAGCCACCGGACATTCACAGCGGTGCTCCACCTGGTTCACAGGCCCCCTGTGCA 60

QY 199 GCGCGCATCGAGGCATTCGCGGAGAGGAGATTTCAGTGTCTTCACGAGGCGCGCGCCGCG 258
DB 61 GCGCGCATCGGGGCATCCGCGGGCCGACTTCAGTGTCTTCACGAGGCGCGGGCCGTG 120

QY 259 GGGCTGGCGGCGACCTTCGCGGCGTTCCTGTGTGTGCGGTGTGACGACCTCTACAGCATC 318
DB 121 GGGCTGGCGGCGACCTTCGCGCGCTTCCTGTCTCGCGCGCTGTGACGACCTGTACAGCATC 180

QY 319 GTGCGCGGCGCCACCGCACCGGGGTGCCGTGTCTAACTCAGGAGCAGGTGCTCTTTC 378
DB 181 GTGCGCGGTGCCACCGCGCAGCGCTGCCCATCGTCAACCTCAAGGACGAGCTGCTGTTT 240

QY 379 CCCAGCTGGAGAGCCTTATTCTCGGGCTCGGAGGCGCAGCTGAAGCCCGGGGCCCGCATC 438
DB 241 CCCAGCTGGAGAGCTCTGTTCTCAGGCTCTGAGGTCCTGTGAAGCCCGGGGCACGCATC 300

QY 439 TTCTCTTTTCACGGCAGAGATGTCTCTGCAGCACCCCGCTGGCCCCCGGAAGAGCGTGTGG 498
DB 301 TTCTCTTTTCACGGCAGGAGCTCTTGAGGCACCCCACTGGCCCCCAGAGAGCGTGTGG 360

QY 499 CACGGCTCGACCCCGACGGGGCGCGGCTGACCGACAGCTACTCGGAGACGTGGCGGACG 558
DB 361 CATGGCTCGGACCCCAACGGGGCGCAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGACG 420

QY 559 GAGGCCCCGCGGCCACCGGGCAGGCGTCTGCTGTGTGGGCGCAGGCTGCTGAGCAG 618
DB 421 GAGGCTCTCTCGGCGCACGGGCCAGGGCTCTCTGCTGTGGGGGCGAGGCTCCTGGGGCAG 480

QY 619 GAGGCGCGGAGCTGCGGCCACCGCTTCGTGGTGTCTTGTCATCGAGAACAGCGTCAATGACC 678
DB 481 AGTGGCGCGAGCTGCCATCACGGCTACATGCTGCTGTGCATTGAGAACAGCTTCATGACT 540

QY 679 TCCTTC 684
DB 541 GCCTCC 546

```

```

RESULT 9
US-10-080-797-2
; Sequence 2, Application US/10080797
; Publication No. US20020183253A1
; GENERAL INFORMATION:
; APPLICANT: Campochiaro, Peter A.
; APPLICANT: Dixon, Katharine H.
; APPLICANT: Brazzell, Romulus K.
; TITLE OF INVENTION: METHOD FOR TREATING OCULAR
; TITLE OF INVENTION: NEOVASCULARIZATION
; FILE REFERENCE: 4-31881A
; CURRENT APPLICATION NUMBER: US/10/080,797
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 551

```

```
; TYPE: DNA
; ORGANISM: Human
US-10-080-797-2

Query Match      50.3%; Score 417; DB 13; Length 551;
Best Local Similarity 85.3%; Pred. No. 5e-89;
Matches 465; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 140 ACACCCACGAGGCTTCCAGCTGGTGTGACCTGGTGGCCCTGAAACAGCCCGCAGCCGG 199
DB 1 ACAGCCACCCGAGCTTCCAGCCGGTGTCTCCACCTGGTGGCTCAACAGCCCGCTGTGAG 60
QY 200 GCAGCATGCGAGGATCCGCGGAGCGGACTTCCAGTGTCTTCCAGCAGGCGCGCGCGCG 259
DB 61 GCGGATGCGGGGATCCGCGGGCGGACTTCCAGTGTCTTCCAGCAGGCGCGCGCGTGG 120
QY 260 GGCTGGCGGACCTTCCGCGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 319
DB 121 GGCTGGCGGACCTTCCGCGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 320 TGGCGCGCGGACCGGACCGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 379
DB 181 TGGCGCGTGGCGACCGGCGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
QY 380 CCAGCTGGAGGCTTATTCTCGGGCTCCGAGGCGGAGCTGAAAGCCCGGGCGCGCATCT 439
DB 241 CCAGCTGGAGGCTTCTGTCTCAGGCTCTGAGGTCGCTGAAAGCCCGGGCGAGCATCT 300
QY 440 TCTCTTTGACGCGAGAGATGTCTGTGACGACCCCGCTGTGGCCCGGAAAGAGGTGTGGC 499
DB 301 TCTCTTTGACGCGAGAGAGTCTGTGAGGACCCCGCATCTGTGGCCCGGAAAGAGGTGTGGC 360
QY 500 ACGGCTCGACCCCGAGCGGGCGGCGCTGACGACGACTACTGCGAGACGTGCGGACGG 559
DB 361 ATGGCTCGGACCCCGACCGGCGGAGGCTGACCGAGAGCTACTGTGAGACGTGCGGACGG 420
QY 560 AGGCGCGCGGCGGACCGGCGGAGCGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 619
DB 421 AGGCTCTCTCGGACCGGCGGAGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
QY 620 AGGCGCGGAGTGTGCGGCGACGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 679
DB 481 GTGCGCGGAGTGTGCTACGCGCTTACATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
QY 680 CTTTC 684
DB 541 CTTCC 545

RESULT 10
US-10-131-241-51
; Sequence 51, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/10131.241
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
; LENGTH: 632
; TYPE: DNA
; ORGANISM: Murinae sp.
```

```
US-10-131-241-51

Query Match      50.1%; Score 415.6; DB 14; Length 632;
Best Local Similarity 82.8%; Pred. No. 1.1e-88;
Matches 523; Conservative 0; Mismatches 29; Indels 80; Gaps 1;

QY 139 CACACCCACGAGGACTTCCAGCTGGTGTGTCACCTGGTGGCCCTGAACAGCCCGCAGCCG 198
DB 1 CATACTCATCAGGACTTTTCAGCCAGTGTCTCCACTGCTGTGCACTGAACACCCCGCTGTCT 60
QY 199 GCGGCGATCGAGGAGCATCCGCGGAGCGGACTTCCAGTGTCTTCCAGCAGGCGCGCGCG 258
DB 61 GGAGGCGATCGTGTGTATCCGCGGAGCAGATTTCCAGTGTCTTCCAGCAAGCCGAGCGCGT 120
QY 259 GGGCTGGCGCGGACCTTCCGCGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 298
DB 121 GGGCTGTGGGACCTTCCGCGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 299 ----- 298

DB 181 GTGCGCGCTGTGACCGGGGCTGTGCCCCTGTCACCTGAAGGACGAGGTGTATCT 240
QY 299 TGAGGACCTCTACAGCATGTGTGCGCGCGCGGACCGGAGTGTGCGGCTGTGTCAGCC 358
DB 241 TGAGGACCTCTACAGCATGTGTGCGCGCGCGGACCGGAGTGTGCGGCTGTGTCAGCC 300
QY 359 TCAGGAGCAGGCTGTCTTCCCGAGCTGGGAGGCGCTTATTCTCGGCTCCGAGGGCCAGC 418
DB 301 TCAGGAGCAGGCTGTCTTCCCGAGCTGGGAGGCGCTTATTCTCGGCTCCGAGGGCCAGC 360
QY 419 TGAAGCCCGGGCGCGCATCTTCTCTTTCGAGCGGACAGAGATGTCTGCGACACCCCGCCT 478
DB 361 TGAAGCCCGGGCGCGCATCTTCTCTTTCGAGCGGACAGAGATGTCTGCGACACCCCGCCT 420
QY 479 GGGCCCGGAGAGCGTGTGCGACCGGCTCCGACCCGAGCGGCGCGCGCTGTGAGCGAGCT 538
DB 421 GGGCCCGGAGAGCGTGTGCGACCGGCTCCGACCCGAGCGGCGCGCGCTGTGAGCGAGCT 480
QY 539 ACTGCGAGACGTGGCGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTGCGCTGG 598
DB 481 ACTGCGAGACGTGGCGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTGCGCTGG 540
QY 599 CCGGCGAGGCTGTGAGGAGGCGCGCGAGCTGCGCGCGCGCGCGCGCGCGCGCTGTGCGCTGTGCA 658
DB 541 CCGGCGAGGCTGTGAGGAGGCGCGCGAGCTGCGCGCGCGCGCGCGCGCGCGCTGTGCGCTGTGCA 600
QY 659 TCAGGAGACGCTCATGACCTCTTCTTCCAAG 690
DB 601 TCAGGAGACGCTCATGACCTCTTCTTCCAAG 632

RESULT 11
US-10-042-347-6
; Sequence 6, Application US/10042347
; Publication No. US20030114370A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Endostatin Protein and Peptide Fragments
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 05213-0880 (43170-249874)
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: US 09/315,689
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 60/106,343
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 09/154,302
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 08/740,168
; PRIOR FILING DATE: 1996-10-22
; PRIOR APPLICATION NUMBER: US 60/005,835
; PRIOR FILING DATE: 1995-10-23
```

;
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-347-6

Query Match 49.6%; Score 410.8; DB 14; Length 534;
Best Local Similarity 85.6%; Pred. No. 1.4e-87;
Matches 457; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
QY 151 GACTTCCAGCTGGTGTGCTGACCTGGGCGCCCTGAAACAGCCCGCGCGGGCGGCATGCCA 210
Db 1 GACTTCCAGCGGTGTCTCCACCTGGTGGCTCAACAGCCCCCTGTACGGCGGCATGCCG 60
QY 211 GGCATCCGGGAGCGGACTTCCAGTGTCTTCCAGCAGCGCGCGCGGGGTGGCGCGC 270
Db 61 GGCATCCGGGCGGACTTCCAGTGTCTTCCAGCAGCGCGCGGGGTGGCGGCGC 120
QY 271 ACCTTCGGGGCTTCTGTGTCGGGTGCGAGGACTCTACAGCATGTGCGCGCGCGC 330
Db 121 ACCTTCGGGGCTTCTGTGTCGGGTGCGAGGACTCTACAGCATGTGCGCGCGTGC 180
QY 331 GACCGCACCGGGTGCCTGTCAACCTCAGGAGGAGGTGCTCTTCCAGCTGGGAG 390
Db 181 GACCGCGACCGGTGCCCCATCGTCAACCTCAAGGACGAGTGTGTTTCCAGCTGGGAG 240
QY 391 GCCTTATTCTCGGGCTCCGAGGGCCAGCTGAAGCCCGGGCCCGCATCTTCTTTTCGAC 450
Db 241 GCTCTGTTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGCCAGCATCTTCTTTTAC 300
QY 451 GGCAGAGATGTCTCTCAGCACCCCGCTGGCCCGGAGAGCGTGTGGCACCGGCTCCGAC 510
Db 301 GGCAGGAGCTCTCTGAGGACACCCACCTGGCCCGGAGAGCGTGTGGCATGGCTCGGAC 360
QY 511 CCCAGGGGCGCGCTGACCGACAGCTACTCGGAGACGTGGCGGACGAGGCGCGGGC 570
Db 361 CCCAACGGGCGCAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGACGAGGCTCCCTCG 420
QY 571 GCCACCGGACGCGTCTGCTGTGGCGGACGAGTGTCTGAGAGCGTGGCGGAGCGCGGAC 630
Db 421 GCCACGGGCGAGGCTCTGCTGTGGGGGCGAGGCTCTGGGGCAGAGTGCCTGGGAGC 480
QY 631 TGCCGCCACGCTTCTGCTGTGCTCTGATCGAGAACAGCGTCTATGACCTCTTC 684
Db 481 TGCCATCAGCCCTACATCGTGTCTGCAATTGAGAACAGCTTCACTGCTCTCC 534

RESULT 12
US-10-131-241-59
; Sequence 59, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1

;
; SEQ ID NO 59
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-241-59
Query Match 49.6%; Score 410.8; DB 14; Length 537;
Best Local Similarity 85.6%; Pred. No. 1.4e-87;
Matches 457; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
QY 151 GACTTCCAGCTGGTGTGCTGACCTGGGCGCCCTGAAACAGCCCGCGCGGGCGGCATGCCA 210
Db 1 GACTTCCAGCGGTGTCTCCACCTGGTGGCTCAACAGCCCCCTGTACGGCGGCATGCCG 60
QY 211 GGCATCCGGGAGCGGACTTCCAGTGTCTTCCAGCAGCGCGCGGGGTGGCGCGC 270
Db 61 GGCATCCGGGCGGACTTCCAGTGTCTTCCAGCAGCGCGGGGTGGCGGCGC 120
QY 271 ACCTTCGGGGCTTCTGTGTCGGGTGCGAGGACTCTACAGCATGTGCGCGCGCGC 330
Db 121 ACCTTCGGGGCTTCTGTGTCGGGTGCGAGGACTCTACAGCATGTGCGCGCGTGC 180
QY 331 GACCGCACCGGGTGCCTGTCAACCTCAGGAGGAGGTGCTCTTCCAGCTGGGAG 390
Db 181 GACCGCGACCGGTGCCCCATCGTCAACCTCAAGGACGAGTGTGTTTCCAGCTGGGAG 240
QY 391 GCCTTATTCTCGGGCTCCGAGGGCCAGCTGAAGCCCGGGCCCGCATCTTCTTTTCGAC 450
Db 241 GCTCTGTTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGCCAGCATCTTCTTTTAC 300
QY 451 GGCAGAGATGTCTCTCAGCACCCCGCTGGCCCGGAGAGCGTGTGGCACCGGCTCCGAC 510
Db 301 GGCAGGAGCTCTCTGAGGACACCCACCTGGCCCGGAGAGCGTGTGGCATGGCTCGGAC 360
QY 511 CCCAGGGGCGCGCTGACCGACAGCTACTCGGAGACGTGGCGGACGAGGCGCGGGC 570
Db 361 CCCAACGGGCGCAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGACGAGGCTCCCTCG 420
QY 571 GCCACCGGACGCGTCTGCTGTGGCGGACGAGTGTCTGAGAGCGTGGCGGAGCGCGGAC 630
Db 421 GCCACGGGCGAGGCTCTGCTGTGGGGGCGAGGCTCTGGGGCAGAGTGCCTGGGAGC 480
QY 631 TGCCGCCACGCTTCTGCTGTGCTCTGATCGAGAACAGCGTCTATGACCTCTTC 684
Db 481 TGCCATCAGCCCTACATCGTGTCTGCAATTGAGAACAGCTTCACTGCTCTCC 534

RESULT 13
US-10-131-241-48
; Sequence 48, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 540
; TYPE: DNA
; ORGANISM: Rhesus monkey
US-10-131-241-48


```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 565 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 37:
US-10-036-869-37

Query Match      45.4%; Score 376.6; DB 13; Length 565;
Best Local Similarity 79.6%; Pred.No.1.5e-79;
Matches 445; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 135 CACCCACACCCACAGGACTTCAGTGTGTGCTGACCTGGTGGCCCTGAACAGCCCGCA 194
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7 CATGCATATCTCATCAGGACTTTTCAGCCAGTGTCTCCACCTGGTGGCACTGAACACCCCT 66

QY 195 GCCGGGGCGCATGCGAGGATCCGGGGAGCGGACTTCCAGTGTCTCCAGCAGCGCGCGC 254
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67 GTCGTGAGGATGCGTGGTATCCGTGGAGCAGATTTCCAGTGTCTCCAGCAAGCCGAGC 126

QY 255 CGCGGGGTGGCGGCGACCTTCGGGCTTCTGTCGTCGGGCTGCGAGGACCTCTACAG 314
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
127 CGTGGGGCTGTCGGGCGACCTTCGGGCTTCTGTCCTCTAGGCTGCAGGATCTCTATAG 186

QY 315 CATGTGCGCGCGCGCGACCGCACCGGGGTGCCCGTCTCAACTCAGGAGCAGGTGCT 374
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
187 CATGTGCGCGCGTGTGACCGGGGTCTGTGCCCATCGTCAACTGAAGGAGGAGGTGCT 246

QY 375 CTTCCCGACGTGGAGGCGCTTATCTCGGGCTCCGAGGGCCAGCTGAAGCCCGGGCCCG 434
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
247 ATCTCCCGACGTGGAGCTCCCGTGTCTGCTCCAGGGTCAACTGAAGCCCGGGCCCG 306

QY 435 CATCTTCTCTTTTCGACGGCAGAGATGCTCTGAGCACCCCGCTGGCCCGGAAAGAGCGT 494
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
307 CATCTTCTTTTGACGGCAGAGATGCTCTGAGCACCCAGCCTGGCCCGCAGAGAGCGT 366

QY 495 GTGGCAGCGCTCGACCCCGCGGGCGCGCTTGACCGACAGCTACTGCGAGACGTGGCG 554
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
367 ATGGCAGCGCTCGACCCCGCGGGCGAGGCTGTATGAGAGTTACTGTGAGACATGGCG 426

QY 555 GACGAGGCGCGCGCGCGCACCGGCGAGCGTCTGCTGGCGGCGAGGCTGCTGGA 614
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
427 AACTGAACTACTGGGGCTACAGGTCAAGGCTCTCTCCCTGTGTAGGAGGCTCTCTGGA 486

QY 615 GCAGGAGGCGCGAGTGGCGCGCACCGCTTCTGTTGGTCTCTGATCGAGAACAGCGTCTAT 674
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
487 ACAGAAAGCTGGAGCTGGCACAACAGCTACATCGTCTGTGATTTGAGATAGCTTCTAT 546

QY 675 GACCTCTTCTCAAGTAG 693
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
547 GACCTCTTCTCAAGTAG 565

```

Search completed: August 17, 2003, 16:34:20
Job time : 321.875 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 13:19:37 ; Search time 2416.32 Seconds
(without alignments)
8338.472 Million cell updates/sec

Title: US-09-938-391-1

Perfect score: 829

Sequence: 1 cctcggcgccagatgacat.....aataaaggagccaaag 829

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_estl:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gssl:*

29: gb_gss2:*

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	493.2	59.5	881	14	CD105862
2	459.6	55.4	929	13	BQ672290 AGENCOURT
3	451	54.4	979	13	BQ673186 AGENCOURT
4	445.2	53.7	551	10	BF074459 221883 MA

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1

CD105862

LOCUS

DEFINITION

AGENCOURT 14021788 NIH MGC 179 Homo sapiens cDNA clone

IMAGE:30355831 5', mRNA sequence.

ACCESSION

CD105862

VERSION

CD105862.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 881)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein

cDNA Library Preparation: Invitrogen Corp

DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov

Plate: NDAM424 row: g column: 24

High quality sequence stop: 689..

CD105862 881 bp mRNA linear EST 15-MAY-2003

AGENCOURT 14021788 NIH MGC 179 Homo sapiens cDNA clone

IMAGE:30355831 5', mRNA sequence.

ACCESSION

CD105862

VERSION

CD105862.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 881)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein

cDNA Library Preparation: Invitrogen Corp

DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov

Plate: NDAM424 row: g column: 24

High quality sequence stop: 689..

5 441.6 53.3 832 10 BG387051 602454749
6 428 51.6 843 9 BF385854 602046021
7 425.8 51.4 715 9 AU125614 602046021
8 425.8 51.4 947 13 BU556872 AGENCOURT
9 422.8 51.0 707 10 BE908201 601500458
10 422.4 51.0 757 10 BE906253 601502237
11 417 50.3 874 12 B1412588 602990468
12 401.6 48.4 657 14 CB444165 695295 MA
13 398.2 48.0 944 13 BU859398 AGENCOURT
14 394.4 47.6 703 13 BU615520 UI-H-FGO-
15 387 46.7 683 12 BM683067 UI-E-E01-
16 386.6 46.6 682 9 AW089583 xk20f03.x
17 379.6 45.8 618 9 AV696242 AGENCOURT
18 379.2 45.7 715 14 CB596713 AGENCOURT
19 378.2 45.6 897 12 B1080524 AGENCOURT
20 376.8 45.5 611 10 AW911243 ur83h10.y
21 376.6 45.4 782 9 AI326391 mm18f09.x
22 376 45.4 733 10 BG072504 H311D11-
23 374.4 45.2 884 12 B1161007 602865213
24 373.8 45.1 753 12 B1904605 603168411
25 373.4 45.0 835 12 B1526580 602925454
26 368.6 44.5 723 12 B1247582 602960041
27 366.6 44.2 1093 13 BQ723254 AGENCOURT
28 356.8 43.0 915 10 BF166139 AGENCOURT
29 356.6 43.0 664 13 BU632049 AGENCOURT
30 354.2 42.7 650 9 AI858615 w140f01.x
31 353.8 42.7 846 13 BU540812 AGENCOURT
32 353 42.6 720 12 B1147444 602914008
33 352.6 42.5 639 12 BM998137 UI-H-DT1-
34 350 42.2 634 13 BQ772348 UI-H-EZ1-
35 350 42.2 660 9 AW192502 UI-H-EZ1-
36 349.4 42.1 652 13 BU352506 603527982
37 348.8 42.1 747 12 BG967333 602833649
38 348.6 42.1 587 9 AA288198 v015b10.x
39 348.2 42.0 634 9 AI970297 w09c02.x
40 343.8 41.5 745 9 AW243446 x095d11.x
41 343.6 41.4 690 12 B1219399 602936756
42 341.8 41.2 556 14 CB216999 NISC nq11
43 341.4 41.2 614 13 BU459935 603367327
44 340.8 41.1 653 10 BF384828 602046324
45 340.4 41.1 623 14 CB424313 598557 MA

Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R., Quackenbush J. and Keele,J.W.
Porcine gene discovery by normalized cDNA-library sequencing and EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
12226715
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCAGCAGC
Plate: 89 row: I column: 13
Seq primer: ATTTAGTGACACTATAG.
Location/Qualifiers
1. .451
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC lPIG"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20, and 30 embryos."

BASE COUNT 75 a 153 c 152 g 71 t
ORIGIN

Alignment Scores:
Pred. No.: 1.25e-30 Length: 451
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19.13% Indels: 0
DB: 10 Gaps: 0

US-09-938-391-2 (1-230) x BG383960 (1-451)

QY 118 ValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGlu 137
|||||
Db 34 GTTAACCTCAGGACGAGGTGCTGTTCCACAGCTGGAGGCTTGTCTCGGGCTCTGAG 93
|||||

QY 138 GlyGlnLeuLysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHis 157
|||||
Db 94 GGCCAGCTGAGCCCGCGCCGCATCTTCTTTTCAGCGGACGACGTCCTTCAGCAC 153
|||||

QY 158 ProAlaTrpPro 161
|||||
Db 154 CTGCTGGGCC 165
|||||

RESULT 2
BG383970
LOCUS
DEFINITION
302444 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION
BG383970
VERSION
BG383970.1 GI:13308442
KEYWORDS
EST.
SOURCE
Sus scrofa (pig)
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 491)
Fahrenkrug,S.C., Smith,T.P.L., Preking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R., Quackenbush J. and Keele,J.W.
Porcine gene discovery by normalized cDNA-library sequencing and

EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
12226715
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCAGCAGC
Plate: 89 row: K column: 13
Seq primer: ATTTAGTGACACTATAG.
Location/Qualifiers
1. .491
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC lPIG"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20, and 30 embryos."

BASE COUNT 82 a 163 c 172 g 74 t
ORIGIN

Alignment Scores:
Pred. No.: 1.37e-30 Length: 491
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19.13% Indels: 0
DB: 10 Gaps: 0

US-09-938-391-2 (1-230) x BG383970 (1-491)

QY 118 ValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGlu 137
|||||
Db 34 GTTAACCTCAGGACGAGGTGCTGTTCCACAGCTGGAGGCTTGTCTCGGGCTCTGAG 93
|||||

QY 138 GlyGlnLeuLysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHis 157
|||||
Db 94 GGCCAGCTGAGCCCGCGCCGCATCTTCTTTTCAGCGGACGACGTCCTTCAGCAC 153
|||||

QY 158 ProAlaTrpPro 161
|||||
Db 154 CTGCTGGGCC 165
|||||

RESULT 3
AW464343
LOCUS
DEFINITION
BP230015B10C11 Scores normalized bovine placenta Bos taurus CDNA
clone BP230015B10C11 5', mRNA sequence.
ACCESSION
AW464343
VERSION
AW464343.1 GI:7034511
KEYWORDS
EST.
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 279)
Lewin,H.A., Soares,M.B., Rebeiz,M., Pardinas,J., Liu,L. and Larson J.H.
Bovine ESTs
Unpublished
Contact: Lewin, H. A.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 18, 2003, 00:28:06 ; Search time 2038.89 Seconds
(without alignments)
2741.704 Million cell updates/sec

Title: US-09-938-391-2
Perfect score: 230
Sequence: 1 PWRADILAGPRLDPPQY.....CRHAFVVLCIENSVMTSFSK 230

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Word size: 1

Total number of hits satisfying chosen parameters: 45562604

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: @
-MODEL=frame_p2n.model -DRV=xlpl
-Q=/cgn2_1/USPTO.spool_p/US09938391/runat_04082003_130739_27598/app_query.fasta_1.718
-DB=EST -QFMT=fastap -SUFFIX=p2noli.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938391_@CGN_1_14399_@runat_04082003_130739_27598 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pbg:*
27: em_gss_vrl:*
28: gb_gss1:

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	44	19.1	451	10	BG383960	BG383960 302432 MA
2	44	19.1	491	10	BG383970	BG383970 302444 MA
3	43	18.7	279	9	AW464343	AW464343 BP230015B
4	43	18.7	551	10	BF074459	BF074459 221883 MA
5	43	18.7	623	14	CB424313	CB424313 598557 MA
6	43	18.7	657	14	CB444165	CB444165 695295 MA
7	43	18.7	709	14	CB443805	CB443805 694911 MA
8	37	16.1	387	10	BF601253	BF601253 266182 MA
9	32	13.9	483	12	BG823096	BG823096 602726433
10	27	11.7	577	12	BM488074	BM488074 pgm2n.pk0
11	27	11.7	614	13	BU459935	BU459935 603367327
12	27	11.7	652	13	BU352506	BU352506 603527982
13	27	11.7	678	14	CD215096	CD215096 pgm2n.pk0
14	27	11.7	686	13	BU290652	BU290652 604164521
15	27	11.7	706	13	BU439577	BU439577 603208252
16	27	11.7	717	13	BU230900	BU230900 603400730
17	27	11.7	736	13	BU268246	BU268246 603507092
18	27	11.7	755	13	BU262659	BU262659 603505164
19	27	11.7	756	13	BU380308	BU380308 603582387
20	27	11.7	776	13	BU485669	BU485669 603472003
21	26	11.3	390	9	AV660284	AV660284 AV660284
22	26	11.3	556	14	CB216999	CB216999 NISC.nq11
23	26	11.3	616	13	BU158380	BU158380 AGENCOURT
24	26	11.3	634	13	BQ772348	BQ772348 UI-H-EZ1-
25	26	11.3	639	12	BM998137	BM998137 UI-H-DT1-
26	26	11.3	650	9	AI858615	AI858615 w140f01.x
27	26	11.3	660	9	AW192502	AW192502 x145e10.x
28	26	11.3	664	13	BU632049	BU632049 UI-H-FE1-
29	26	11.3	682	9	AW089583	AW089583 x320f03.x
30	26	11.3	683	12	BM683067	BM683067 UI-E-E01-
31	26	11.3	703	13	BU615520	BU615520 UI-H-FGO-
32	26	11.3	707	10	BE908201	BE908201 601500458
33	26	11.3	715	9	AU125614	AU125614 AU125614
34	26	11.3	757	10	BE906253	BE906253 601502237
35	26	11.3	832	10	BG387051	BG387051 602454749
36	26	11.3	846	13	BU540812	BU540812 AGENCOURT
37	26	11.3	881	14	CD105862	CD105862 AGENCOURT
38	26	11.3	884	12	BI161007	BI161007 602865213
39	26	11.3	929	13	BQ672290	BQ672290 AGENCOURT
40	26	11.3	944	13	BU859398	BU859398 AGENCOURT
41	26	11.3	947	13	BU556872	BU556872 AGENCOURT
42	26	11.3	979	13	BQ673186	BQ673186 AGENCOURT
43	26	11.3	1093	13	BQ723254	BQ723254 AGENCOURT
44	24	10.4	409	14	CB537919	CB537919 775852 MA
45	23	10.0	302	10	BF117397	BF117397 uz36e11.y

ALIGNMENTS

RESULT 1
BG383960 451 bp mRNA linear EST 12-MAR-2001
LOCUS 302432 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
DEFINITION BG383960
ACCESSION BG383960
VERSION BG383960.1 GI:13308432
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 451)
AUTHORS Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,

Qy 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 1624 GGGCTGGGGGACCTTCGGCGCTTCCTGCTCGCGCTGCAGGACCTGTACAGCATC 1683

Qy 107 ValArgArgAlaAspArg 112
Db 1684 GTGCGCGGTGCCGACCGC 1701

RESULT 13

US-10-060-036-144
; Sequence 144, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144
; LENGTH: 4551
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-036-144

Alignment Scores:
Pred. No.: 8 65e-16 Length: 4551
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-2 (1-230) x US-10-060-036-144 (1-4551)

Qy 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 4120 GGGCTGGGGGACCTTCGGCGCTTCCTGCTCGCGCTGCAGGACCTGTACAGCATC 4179

Qy 107 ValArgArgAlaAspArg 112
Db 4180 GTGCGCGGTGCCGACCGC 4197

RESULT 14

US-10-131-241-48
; Sequence 48, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 540
; TYPE: DNA

; ORGANISM: Rhesus monkey
US-10-131-241-48

Alignment Scores:
Pred. No.: 1.32e-13 Length: 540
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.00% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-2 (1-230) x US-10-131-241-48 (1-540)

Qy 90 GlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIleValArgArg 109
Db 130 GGCACCTTCGGGTGCTTCCTGTCTCAGGCTGCAGGACCTGTACAGCATCGTGCGCCGT 189

Qy 110 AlaAspArg 112
Db 190 GCCGACCGC 198

RESULT 15

US-10-292-418-17
; Sequence 17, Application US/10292418
; Publication No. US20030139365A1
; GENERAL INFORMATION:
; APPLICANT: Lo, Kin-Ming
; APPLICANT: Li, Yue
; APPLICANT: Gillies, Stephen D
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
; TITLE OF INVENTION: Immunofusins
; FILE REFERENCE: LEX-006C1
; CURRENT APPLICATION NUMBER: US/10/292,418
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/383,315
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: US 60/097,883
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(552)
; OTHER INFORMATION: endostatin
US-10-292-418-17

Alignment Scores:
Pred. No.: 1.32e-13 Length: 552
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.00% Indels: 0
DB: 12 Gaps: 0

US-09-938-391-2 (1-230) x US-10-292-418-17 (1-552)

Qy 90 GlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIleValArgArg 109
Db 130 GGCACCTTCGGGTGCTTCCTGTCTCAGGCTGCAGGACCTGTATAGCATCGTGCGCCGT 189

Qy 110 AlaAspArg 112
Db 190 GCTGACCGC 198

Search completed: August 18, 2003, 05:48:57
Job time : 917.167 secs

; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Human
US-10-080-797-2

Alignment Scores:
Pred. No.: 1,438-16 Length: 551
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 13 Gaps: 0

US-09-938-391-2 (1-230) x US-10-080-797-2 (1-551)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
DB 120 GGGCTGGCGGGACCTTCGGCGCTTCTGTCTCGCGCTGCAGGACCTGTACAGCATC 179
QY 107 ValArgArgAlaAspArg 112
DB 180 GTGCGCGGTGGCGACGC 197

RESULT 10

US-09-873-676-30

; Sequence 30, Application US/09873676
; Patent No. US2002007289A1
; GENERAL INFORMATION:
; APPLICANT: MacDonald, Nicholas J.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Argiostatin and Endostatin Binding Proteins and Methods of Use
; FILE REFERENCE: 05213-0378 (43170-259333)
; CURRENT APPLICATION NUMBER: US/09/873,676
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-676-30

Alignment Scores:
Pred. No.: 1,438-16 Length: 552
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 9 Gaps: 0

US-09-938-391-2 (1-230) x US-09-873-676-30 (1-552)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
DB 121 GGGCTGGCGGGACCTTCGGCGCTTCTGTCTCGCGCTGCAGGACCTGTACAGCATC 180
QY 107 ValArgArgAlaAspArg 112
DB 181 GTGCGCGGTGGCGACGC 198

RESULT 11

US-10-060-036-64

; Sequence 64, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.

; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yudi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-036-64

Alignment Scores:
Pred. No.: 1,488-16 Length: 574
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-2 (1-230) x US-10-060-036-64 (1-574)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
DB 455 GGGCTGGCGGGACCTTCGGCGCTTCTGTCTCGCGCTGCAGGACCTGTACAGCATC 514
QY 107 ValArgArgAlaAspArg 112
DB 515 GTGCGCGGTGGCGACGC 532

RESULT 12

US-09-880-107-2178
; Sequence 2178, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2178
; LENGTH: 3394
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L22548
US-09-880-107-2178

Alignment Scores:
Pred. No.: 6,748-16 Length: 3394
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 10 Gaps: 0

US-09-938-391-2 (1-230) x US-09-880-107-2178 (1-3394)


```
; CURRENT APPLICATION NUMBER: US/10/042,347
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: US 09/315,689
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 60/106,343
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 09/154,302
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 08/740,168
; PRIOR FILING DATE: 1996-10-22
; PRIOR APPLICATION NUMBER: US 60/005,835
; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-347-4

Alignment Scores:
Pred. No.: 1,42e-16 Length: 546
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-2 (1-230) x US-10-042-347-4 (1-546)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 121 GGGCTGGCGGACCTTCGGCGCTTCCTGCTCGCGGCTGCGAGCCTGTACAGCATC 180

QY 107 ValArgArgAlaAspArg 112
Db 181 GTGCGCGCGTGGCCGACCGC 198

RESULT 7
US-10-292-418-3
; Sequence 3, Application US/10292418
; Publication No. US20030139365A1
; GENERAL INFORMATION:
; APPLICANT: Lo, Kin-Ming
; APPLICANT: Li, Yue
; APPLICANT: Gillies, Stephen D
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
; FILE OF INVENTION: Immunofusins
; FILE REFERENCE: LEX-006C1
; CURRENT APPLICATION NUMBER: US/10/292,418
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/383,315
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: US 60/097,883
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(549)
; OTHER INFORMATION: endostatin
US-10-292-418-3

Alignment Scores:
Pred. No.: 1,42e-16 Length: 549
```

```
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 12 Gaps: 0

US-09-938-391-2 (1-230) x US-10-292-418-3 (1-549)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 121 GGGCTGGCGGACCTTCGGCGCTTCCTGCTCGCGGCTGCGAGCCTGTACAGCATC 180

QY 107 ValArgArgAlaAspArg 112
Db 181 GTGCGCGCGTGGCCGACCGC 198

RESULT 8
US-10-131-241-53
; Sequence 53, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-241-53

Alignment Scores:
Pred. No.: 1,42e-16 Length: 549
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-2 (1-230) x US-10-131-241-53 (1-549)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 121 GGGCTGGCGGACCTTCGGCGCTTCCTGCTCGCGGCTGCGAGCCTGTACAGCATC 180

QY 107 ValArgArgAlaAspArg 112
Db 181 GTGCGCGCGTGGCCGACCGC 198

RESULT 9
US-10-080-797-2
; Sequence 2, Application US/10080797
; Publication No. US20020183253A1
; GENERAL INFORMATION:
; APPLICANT: Campocharo, Peter A.
; APPLICANT: Dixon, Katharine H.
; APPLICANT: Brazzell, Romulus K.
; TITLE OF INVENTION: METHOD FOR TREATING OCULAR
; FILE REFERENCE: 4-31881A
; CURRENT APPLICATION NUMBER: US/10/080,797
; CURRENT FILING DATE: 2002-02-21
```

US-10-131-241-51

Alignment Scores:

Pred. No.: 2 45e-120 Length: 632
Score: 131.00 Matches: 131
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 56.96% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-2 (1-230) x US-10-131-241-51 (1-632)

QY 100 LeuGlnAspLeuTyrSerIleValArgAlaAspArgThrGlyValProValValAsn 119
DB 240 TTGCAGGACCTCTACAGCATCGTGCGCGCGCGACCGACCGGGGTGCCGTGCTCAAC 299
QY 120 LeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGln 139
DB 300 CTCAGGGACGAGGTGCTCTTCCACGCTGGAGGCGCTTATCTCGGGCTCCGAGGGCCAG 359
QY 140 LeuLysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAla 159
DB 360 CTGAACCCCGGGCCCGCATCTTCTTCGACGGCAGAGATGCTCTGACGACACCCCGCC 419
QY 160 TrpProArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSer 179
DB 420 TGGCCCCGGAAGCGGTGTGGACGCTCCGACCCCGCGGGCGCGCTGACCGACAGC 479
QY 180 TyrCysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeu 199
DB 480 TACTGCGAGCGTGGCGGAGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTG 539
QY 200 AlaGlyArgLeuLeuGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCys 219
DB 540 GCGGGCAGGCTGTGAGCAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTG 599
QY 220 IleGluAsnSerValMetThrSerPheSerLys 230
DB 600 ATCGACAACAGCGTCATGACCTCTCTTCTCAAG 632

RESULT 4

US-10-042-347-6
; Sequence 6, Application US/10042347
; Publication No. US20030114370A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Endostatin Protein and Peptide Fu
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 05213-0880 (43170-249874)
; CURRENT APPLICATION NUMBER: US/10/042,347
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: US 09/315,689
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 60/106,343
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 09/154,302
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 08/740,168
; PRIOR FILING DATE: 1996-10-22
; PRIOR APPLICATION NUMBER: US 60/005,835
; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-347-6

Alignment Scores:

Pred. No.: 1.39e-16 Length: 534
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-2 (1-230) x US-10-042-347-6 (1-534)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
DB 109 GGGCTGGCGGGCACCTTCGCGCCCTTCTGTCTCGCGCTCGAGGACCTGTACAGCATC 168
QY 107 ValArgArgAlaAspArg 112
DB 169 GTGCGCGGTGCGGACCGC 186

RESULT 5

US-10-131-241-59
; Sequence 59, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifera
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271585
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-241-59

Alignment Scores:

Pred. No.: 1.4e-16 Length: 537
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-2 (1-230) x US-10-131-241-59 (1-537)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
DB 109 GGGCTGGCGGGCACCTTCGCGCCCTTCTGTCTCGCGCTCGAGGACCTGTACAGCATC 168
QY 107 ValArgArgAlaAspArg 112
DB 169 GTGCGCGGTGCGGACCGC 186

RESULT 6

US-10-042-347-4
; Sequence 4, Application US/10042347
; Publication No. US20030114370A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Endostatin Protein and Peptide Pri
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 05213-0880 (43170-249874)

LOCATION: (1)..(552)
OTHER INFORMATION: Endostatin
US-10-292-418-34

Alignment Scores: 7,01e-165 Length: 552
Pred. No.: 176.00 Matches: 176
Score: 176.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 76.52% Indels: 0
DB: 12 Gaps: 0

US-09-938-391-2 (1-230) x US-10-292-418-34 (1-552)

QY 55 ValLeuHisLeuValAlaLeuAenSerProGlnProGlyGlyMetArgGlyLeArgGly 74
DB 25 GTGCTGCACCTGGTGGCCCTGAACAGCCGCGCGGCGCATCGAGGCGATCCGGGA 84
QY 75 AlaAspPheGlnCysPheGlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAla 94
DB 85 GCGGACTTCCAGTGTCTCCAGCAGGCGCGCGCGGCGTGGCGGACCTTCGGGGCC 144
QY 95 PheLeuSerSerArgLeuGlnAspLeuTyrSerIleValArgAlaAspArgThrGly 114
DB 145 TTCCTGTGTGGCGGTGCAGGACCTCTACAGCATCGTGGCGCGCGCGCACCGCGG 204
QY 115 ValProValValLeuAspLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSer 134
DB 205 GTCCCGTGTGTAACCTCAGGAGCAGAGGTGCTTCCCGAGCTGGGAGGCTTATTCG 264
QY 135 GlySerGluGlyGlnLeuLysProGlyAlaArgIlePheSerPheAspGlyArgVal 154
DB 265 GGCTCCGAGGCGCAGCTGAAGCCCGGCGCGCGCATCTCTTTCGAGCGAGAGATGC 324
QY 155 LeuGlnHisProAlaTrpProArgLysSerValTrpHisGlySerAspProSerGlyArg 174
DB 325 CTGACGACCCCGCTGGCGCGGAGAGCGTGTGGCAGCGGTCTCCGACCCCGCGGCGC 384
QY 175 ArgLeuThrAspSerTyrCysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGln 194
DB 385 CGCTGACCGACAGCTACTCGAGAGCTGGCGGAGGCGCGCGCGCGCGCGCGAG 444
QY 195 AlaSerSerLeuLeuAlaGlyArgLeuGluGlnGlnAlaAlaSerCysArgHisAla 214
DB 445 GCGTGTGTGTCTGGCGGCGAGGCTGTGGAGCAGGAGCGCGGAGTGTGGCGGCGC 504
QY 215 PheValValLeuCysIleGluAenSerValMetThrSerPheSerLys 230
DB 505 TTCGTGTGTCTGTGCATCGAGAACAGCGTGTATGACCTCTCTTCCAAAG 552

RESULT 2

US-10-131-241-51

Sequence 50, Application US/10131241

Publication No. US20030012792A1

GENERAL INFORMATION:

APPLICANT: Fortier, John W.

TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation

FILE REFERENCE: 05213-0344 43170-271565

CURRENT APPLICATION NUMBER: US/10/131,241

CURRENT FILING DATE: 2002-07-22

PRIOR APPLICATION NUMBER: US 09/413,049

PRIOR FILING DATE: 1999-10-06

PRIOR APPLICATION NUMBER: US 09/316,802

PRIOR FILING DATE: 1999-05-21

PRIOR APPLICATION NUMBER: US 60/086,586

PRIOR FILING DATE: 1998-05-22

NUMBER OF SEQ ID NOS: 65

SOFTWARE: PatentIn version 3.1

SEQ ID NO 50

LENGTH: 552

TYPE: DNA

ORGANISM: Murinae sp.

ORGANISM: Canine sp.
US-10-131-241-50

Alignment Scores: 7,01e-165 Length: 552
Pred. No.: 176.00 Matches: 176
Score: 176.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 76.52% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-2 (1-230) x US-10-131-241-50 (1-552)

QY 55 ValLeuHisLeuValAlaLeuAenSerProGlnProGlyGlyMetArgGlyLeArgGly 74
DB 25 GTGCTGCACCTGGTGGCCCTGAACAGCCGCGCGGCGCATCGAGGCGATCCGGGA 84
QY 75 AlaAspPheGlnCysPheGlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAla 94
DB 85 GCGGACTTCCAGTGTCTCCAGCAGGCGCGCGCGGCGTGGCGGACCTTCGGGGCC 144
QY 95 PheLeuSerSerArgLeuGlnAspLeuTyrSerIleValArgAlaAspArgThrGly 114
DB 145 TTCCTGTGTGGCGGTGCAGGACCTCTACAGCATCGTGGCGCGCGCGCACCGCGG 204
QY 115 ValProValValLeuAspLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSer 134
DB 205 GTGCGCGTGTCAACCTCAGGAGCAGAGGTGCTTTCGCCAGCTGGGAGGCTTATTCG 264
QY 135 GlySerGluGlyGlnLeuLysProGlyAlaArgIlePheSerPheAspGlyArgVal 154
DB 265 GGCTCCGAGGCGCAGCTGAAGCCCGGCGCGCGCATCTCTTTCGAGCGAGAGATGC 324
QY 155 LeuGlnHisProAlaTrpProArgLysSerValTrpHisGlySerAspProSerGlyArg 174
DB 325 CTGACGACCCCGCTGGCGCGGAGAGCGTGTGGCAGCGGTCTCCGACCCCGCGGCGC 384
QY 175 ArgLeuThrAspSerTyrCysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGln 194
DB 385 CGCTGACCGACAGCTACTCGAGAGCTGGCGGAGGCGCGCGCGCGCGCGAG 444
QY 195 AlaSerSerLeuLeuAlaGlyArgLeuGluGlnGlnAlaAlaSerCysArgHisAla 214
DB 445 GCGTGTGTGTCTGGCGGCGAGGCTGTGGAGCAGGAGCGCGGAGTGTGGCGGCGC 504
QY 215 PheValValLeuCysIleGluAenSerValMetThrSerPheSerLys 230
DB 505 TTCGTGTGTCTGTGCATCGAGAACAGCGTGTATGACCTCTCTTCCAAAG 552

RESULT 3

US-10-131-241-51

Sequence 51, Application US/10131241

Publication No. US20030012792A1

GENERAL INFORMATION:

APPLICANT: Fortier, John W.

TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation

FILE REFERENCE: 05213-0344 43170-271565

CURRENT APPLICATION NUMBER: US/10/131,241

CURRENT FILING DATE: 2002-07-22

PRIOR APPLICATION NUMBER: US 09/413,049

PRIOR FILING DATE: 1999-10-06

PRIOR APPLICATION NUMBER: US 09/316,802

PRIOR FILING DATE: 1999-05-21

PRIOR APPLICATION NUMBER: US 60/086,586

PRIOR FILING DATE: 1998-05-22

NUMBER OF SEQ ID NOS: 65

SOFTWARE: PatentIn version 3.1

SEQ ID NO 51

LENGTH: 632

TYPE: DNA

ORGANISM: Murinae sp.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 18, 2003, 02:40:12 ; Search time 906.667 Seconds
(without alignments)
567.713 Million cell updates/sec

Title: US-09-938-391-2
Perfect score: 230
Sequence: 1 PWRADDILAGPRLLDPPQY.....CRHAFVVLCSNVTFSFK 230

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1504479 seqs, 1118970152 residues

Word size: 1
Total number of hits satisfying chosen parameters: 3005078

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame.p2n.model -DRV=xlsp
-Q=/cgn2_1/USPTO.spool_p/US0938391/runat_04082003_130741_27700/app_query.fasta_1.718
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2noli.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US0938391 -CGM_1_129_orunat_04082003_130741_27700
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NSG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARG TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FCAPOP=6 -FCAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	176	76.5	552	12	US-10-292-418-34
					Sequence 34, Appl

2	176	76.5	552	14	US-10-131-241-50	Sequence 50, Appl
3	131	57.0	632	14	US-10-131-241-51	Sequence 51, Appl
4	26	11.3	534	14	US-10-042-347-6	Sequence 6, Appl
5	26	11.3	537	14	US-10-131-241-59	Sequence 59, Appl
6	26	11.3	546	14	US-10-042-347-4	Sequence 4, Appl
7	26	11.3	549	12	US-10-292-418-3	Sequence 3, Appl
8	26	11.3	549	14	US-10-131-241-53	Sequence 53, Appl
9	26	11.3	551	13	US-10-080-797-2	Sequence 2, Appl
10	26	11.3	552	9	US-09-873-676-30	Sequence 30, Appl
11	26	11.3	574	14	US-10-080-036-84	Sequence 64, Appl
12	26	11.3	3394	10	US-09-880-107-2178	Sequence 2178, Ap
13	26	11.3	4551	14	US-10-060-036-144	Sequence 144, App
14	23	10.0	540	14	US-10-131-241-48	Sequence 48, Appl
15	23	10.0	552	12	US-10-292-418-17	Sequence 17, Appl
16	23	10.0	558	9	US-09-775-174-3	Sequence 3, Appl
17	23	10.0	558	9	US-09-775-325-3	Sequence 3, Appl
18	23	10.0	565	13	US-10-036-869-37	Sequence 37, Appl
19	23	10.0	573	10	US-09-998-831-12	Sequence 12, Appl
20	23	10.0	624	13	US-10-080-797-4	Sequence 4, Appl
21	20	8.7	900	14	US-10-131-241-45	Sequence 45, Appl
22	13	5.7	434	14	US-10-060-036-4	Sequence 4, Appl
23	10	4.3	287	14	US-10-080-036-97	Sequence 97, Appl
24	9	3.9	96	13	US-10-080-797-14	Sequence 14, Appl
25	9	3.9	540	13	US-10-027-632-186155	Sequence 186155,
26	9	3.9	540	13	US-10-027-632-186156	Sequence 186156,
27	9	3.9	542	13	US-10-027-632-285764	Sequence 285764,
28	9	3.9	601	14	US-10-207-951-24	Sequence 24, Appl
29	9	3.9	652	13	US-10-027-632-264113	Sequence 264113,
30	9	3.9	655	13	US-10-027-632-221236	Sequence 221236,
31	9	3.9	720	13	US-10-027-632-21145	Sequence 21145, A
32	9	3.9	733	13	US-10-027-632-262492	Sequence 262492,
33	9	3.9	740	13	US-10-027-632-147953	Sequence 147953,
34	9	3.9	740	13	US-10-027-632-147954	Sequence 147954,
35	9	3.9	1197	13	US-10-027-632-124065	Sequence 124065,
36	9	3.9	1823	14	US-10-037-270-1012	Sequence 1012, Ap
37	9	3.9	2308	10	US-09-927-267-2	Sequence 2, Appl
38	9	3.9	2589	10	US-03-778-971-4	Sequence 4, Appl
39	9	3.9	3337	12	US-10-024-298A-169	Sequence 169, App
40	9	3.9	10989	9	US-09-735-932-3	Sequence 3, Appl
41	9	3.9	10989	14	US-10-207-951-3	Sequence 3, Appl
42	9	3.9	32221	10	US-09-764-847-1406	Sequence 1406, Ap
43	9	3.9	32221	14	US-10-092-154-1406	Sequence 1406, Ap
44	9	3.9	90541	9	US-09-759-359A-3	Sequence 3, Appl
45	9	3.9	3186778	13	US-10-027-632-174961	Sequence 174961,

ALIGNMENTS

RESULT 1
US-10-292-418-34
; Sequence 34, Application US/10292418
; Publication No. US20030139365A1
; GENERAL INFORMATION:
; APPLICANT: Lo, Kin-Ming
; APPLICANT: Li, Yue
; APPLICANT: Gillies, Stephen D
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
; FILE REFERENCE: Immunofusins
; CURRENT APPLICATION NUMBER: US/10/292,418
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/383,315
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: US 60/097,883
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS

```

; GENERAL INFORMATION:
; APPLICANT: CARRAWAY, KERMIT L.
; APPLICANT: CARRAWAY, CARRAWAY, CORALIE A.
; APPLICANT: FREGIEN, NEVIS L.
; TITLE OF INVENTION: ONCOGENE PRODUCT LIGAND
; NUMBER OF SEQUENCES: 125
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/179,481
; FILING DATE: 28-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/922,521
; FILING DATE: 30-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 200702/UM92-08CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2603 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-179-481-1

Alignment Scores:
Pred. No.: 28.8 Length: 2603
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 1 Gaps: 0

US-09-938-391-2 (1-230) x US-08-179-481-1 (1-2603)

Qy 42 ValHisThrHisThrHisGln 50
Db 2482 GTACACACACACACACACACACACACACAG 2508

RESULT 15
US-09-759-359A-3/c
; Sequence 3, Application US/09759359A
; Patent No. 6492153
; GENERAL INFORMATION:
; APPLICANT: ABU-THREIDEH, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001043
; CURRENT FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 90541
; TYPE: DNA

; ORGANISM: Human
; US-09-759-359A-3

Alignment Scores:
Pred. No.: 760 Length: 90541
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 4 Gaps: 0

US-09-938-391-2 (1-230) x US-09-759-359A-3 (1-90541)

Qy 42 ValHisThrHisThrHisGln 50
Db 37088 GTACACACACACACACACACACACACACAA 37062

Search completed: August 18, 2003, 03:43:58
Job time : 89.6667 secs
```

FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; NAME/KEY: CDS
; LOCATION: (1)..(573)
US-09-561-499-12

Alignment Scores:
Pred. No.: 2.57e-13 Length: 573
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.00% Indels: 0
DB: 4 Gaps: 0

US-09-938-391-2 (1-230) x US-09-561-499-12 (1-573)

QY 90 GlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIleValArgArg 109
DB 151 GGCACCTTCGGGCTTCTCTAGGCTCAGGATCTCTATAGCATCGTGCGCGT 210
QY 110 AlaAspArg 112
DB 211 GCTGACCGG 219

RESULT 12
US-08-159-784-1
; Sequence 1, Application US/08159784
; Patent No. 5643783
; GENERAL INFORMATION:
; APPLICANT: Bjorn R. Olsen
; TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,784
; FILING DATE: December 1, 1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: John F. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00246/170001
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4031
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-159-784-1

Alignment Scores:
Pred. No.: 1.55e-12 Length: 4031
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 10.00% Indels: 0
DB: 1 Gaps: 0

US-09-938-391-2 (1-230) x US-08-159-784-1 (1-4031)

QY 90 GlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIleValArgArg 109
DB 3442 GGCACCTTCGGGCTTCTCTAGGCTCAGGATCTCTATAGCATCGTGCGCGT 3501
QY 110 AlaAspArg 112
DB 3502 GCTGACCGG 3510

RESULT 13

US-09-620-312D-1012/c
; Sequence 1012, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yuning
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 1012
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (164)..(1786)
US-09-620-312D-1012

Alignment Scores:
Pred. No.: 20.7 Length: 1823
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 4 Gaps: 0

US-09-938-391-2 (1-230) x US-09-620-312D-1012 (1-1823)

QY 86 AlaGlyLeuAlaGlyThrPheArgAla 94
DB 1113 GCTGGGTTGGCAGCACCTTTCAGGGCC 1087

RESULT 14
US-08-179-481-1
; Sequence 1, Application US/08179481
; Patent No. 5624816

; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002500
; CURRENT APPLICATION NUMBER: US/09/561,500
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; NAME/KEY: CDS
; LOCATION: (1)..(573)
US-09-561-500-12
Alignment Scores:
Pred. No.: 2,57e-13 Length: 573
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.00% Indels: 0
Gaps: 4
DB: 4

US-09-938-391-2 (1-230) x US-09-561-500-12 (1-573)
Qy 90 GlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyfSerileValArg 109
Db 151 GGCACCTTCGGGCTTCTCTAGGCTGCAGGATCTCTATAGCATCGTGGCCCGT 210
Qy 110 AlaAspArg 112
Db 211 GCTGACCGG 219

RESULT 9
US-09-561-108-12
; Sequence 12, Application US/09561108
; Patent No. 6342221
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002584
; CURRENT APPLICATION NUMBER: US/09/561,108
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; NAME/KEY: CDS
; LOCATION: (1)..(573)
US-09-561-108-12
Alignment Scores:
Pred. No.: 2,57e-13 Length: 573
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.00% Indels: 0
Gaps: 4
DB: 4

US-09-938-391-2 (1-230) x US-09-561-108-12 (1-573)
Qy 90 GlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyfSerileValArg 109
Db 151 GGCACCTTCGGGCTTCTCTAGGCTGCAGGATCTCTATAGCATCGTGGCCCGT 210
Qy 110 AlaAspArg 112
Db 211 GCTGACCGG 219

RESULT 10
US-09-561-526-12
; Sequence 12, Application US/09561526
; Patent No. 6416758
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE KITS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002586
; CURRENT APPLICATION NUMBER: US/09/561,526
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; NAME/KEY: CDS
; LOCATION: (1)..(573)
US-09-561-526-12
Alignment Scores:
Pred. No.: 2,57e-13 Length: 573
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.00% Indels: 0
Gaps: 4
DB: 4

US-09-938-391-2 (1-230) x US-09-561-526-12 (1-573)
Qy 90 GlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyfSerileValArg 109
Db 151 GGCACCTTCGGGCTTCTCTAGGCTGCAGGATCTCTATAGCATCGTGGCCCGT 210
Qy 110 AlaAspArg 112
Db 211 GCTGACCGG 219

RESULT 11
US-09-561-499-12
; Sequence 12, Application US/09561499
; Patent No. 6524583
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002582
; CURRENT APPLICATION NUMBER: US/09/561,499
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Artificial Sequence

RESULT 2
US-09-315-689-4
; Sequence 4, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Folkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-315-689-4
Alignment Scores:
Pred. No.: 3-23e-16 Length: 546
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 4 Gaps: 0
US-09-938-391-2 (1-230) x US-09-315-689-4 (1-546)
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 121 GGCGTGGCGGACCTTCCTGCTCTCGCGCTCGAGCAGCTGTACAGCATC 180
QY 107 ValArgAlaAspArg 112
Db 181 GTGCGCGGTGCGGACCGC 198
RESULT 3
US-09-206-059-30
; Sequence 30, Application US/09206059
; Patent No. 6201104
; GENERAL INFORMATION:
; APPLICANT: MacDonald, Nicholas
; APPLICANT: Sim, Kim Lee
; TITLE OF INVENTION: Angiogenesis-Inhibiting Protein Binding Peptides and
; FILE OF INVENTION: Proteins and Methods of Use
; FILE REFERENCE: 05213-0370
; CURRENT APPLICATION NUMBER: US/09/206,059
; CURRENT FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-206-059-30
Alignment Scores:
Pred. No.: 3-26e-16 Length: 552
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 3 Gaps: 0
US-09-938-391-2 (1-230) x US-09-206-059-30 (1-552)
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 121 GGCGTGGCGGACCTTCCTGCTCTCGCGCTCGAGCAGCTGTACAGCATC 180
QY 107 ValArgAlaAspArg 112

Db 181 GTGCGCGGTGCGGACCGC 198
RESULT 4
US-08-159-784-4
; Sequence 4, Application US/08159784
; Patent No. 5643783
; GENERAL INFORMATION:
; APPLICANT: Bjorn R. Olsen
; TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,784
; FILING DATE: December 1, 1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: John F. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00246/170001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3394
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-159-784-4
Alignment Scores:
Pred. No.: 1-74e-15 Length: 3394
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 1 Gaps: 0
US-09-938-391-2 (1-230) x US-08-159-784-4 (1-3394)
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 1624 GGCGTGGCGGACCTTCCTGCTCTCGCGCTCGAGCAGCTGTACAGCATC 1683
QY 107 ValArgAlaAspArg 112
Db 1684 GTGCGCGGTGCGGACCGC 1701
RESULT 5
US-09-449-293-3
; Sequence 3, Application US/09449293
; Patent No. 6267954
; GENERAL INFORMATION:
; APPLICANT: Abitbol, Marc
; APPLICANT: Uteza, Yves
; APPLICANT: Menasche, Maurice
; APPLICANT: Bossard, Carine

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 18, 2003, 00:40:31 ; Search time 71.6667 Seconds
(without alignments)
1416.532 Million cell updates/sec

Title: US-09-938-391-2
Perfect score: 230
Sequence: 1 PWRADDILGPPRLDPOPY.....CRHAFVWLCIENSVMTSFSK 230

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues
Word size: 1

Total number of hits satisfying chosen parameters: 1135299

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlsp
-Q=/cgn2_1/USPTO.spool_p/US09938391/runat_04082003_130739_27621/app_query.fasta_1.718
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2noli.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938391@cgn_1_133@runat_04082003_130739_27621 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

- 1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/prodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/prodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	11.3	534	US-09-315-689-6	Sequence 6, Appli
2	26	11.3	546	US-09-315-689-4	Sequence 4, Appli
3	26	11.3	552	US-09-206-059-30	Sequence 30, Appli
4	26	11.3	3394	US-08-159-784-4	Sequence 4, Appli
5	23	10.0	558	US-09-449-293-3	Sequence 3, Appli
6	23	10.0	558	US-09-775-325-3	Sequence 3, Appli
7	23	10.0	565	US-08-985-526-37	Sequence 37, Appli
8	23	10.0	573	US-09-561-500-12	Sequence 12, Appli
9	23	10.0	573	US-09-561-108-12	Sequence 12, Appli
10	23	10.0	573	US-09-561-526-12	Sequence 12, Appli
11	23	10.0	573	US-09-561-499-12	Sequence 12, Appli
12	23	10.0	4031	US-08-159-784-1	Sequence 1, Appli

c 13	9	3.9	1823	4	US-09-620-312D-1012	Sequence 1012, Ap
c 14	9	3.9	2603	1	US-08-179-481-1	Sequence 1, Appli
c 15	9	3.9	90541	4	US-09-759-359A-3	Sequence 3, Appli
c 16	8	3.5	132	4	US-09-890-848-7	Sequence 7, Appli
c 17	8	3.5	141	3	US-08-750-064-3	Sequence 3, Appli
c 18	8	3.5	141	3	US-08-545-196B-16	Sequence 16, Appli
c 19	8	3.5	198	4	US-09-918-686-22	Sequence 22, Appli
c 20	8	3.5	202	1	US-08-222-177A-38	Sequence 38, Appli
c 21	8	3.5	208	1	US-08-222-177A-41	Sequence 41, Appli
c 22	8	3.5	244	4	US-09-397-787-58	Sequence 58, Appli
c 23	8	3.5	249	1	US-08-222-177A-28	Sequence 28, Appli
c 24	8	3.5	279	4	US-09-438-906-12	Sequence 12, Appli
c 25	8	3.5	287	4	US-09-544-618-13	Sequence 13, Appli
c 26	8	3.5	323	1	US-08-222-177A-31	Sequence 31, Appli
c 27	8	3.5	381	4	US-09-702-705-1782	Sequence 1782, Ap
c 28	8	3.5	381	4	US-09-736-457-1782	Sequence 1782, Ap
c 29	8	3.5	443	3	US-08-903-139B-17	Sequence 17, Appli
c 30	8	3.5	495	4	US-09-252-991A-8690	Sequence 8690, Ap
c 31	8	3.5	502	1	US-08-384-489-4	Sequence 4, Appli
c 32	8	3.5	508	1	US-08-318-193-69	Sequence 69, Appli
c 33	8	3.5	538	4	US-09-495-050A-105	Sequence 105, App
c 34	8	3.5	576	4	US-09-252-991A-8794	Sequence 8794, Ap
c 35	8	3.5	800	4	US-09-222-938A-66	Sequence 68, Appli
c 36	8	3.5	884	4	US-09-171-209-76	Sequence 76, Appli
c 37	8	3.5	965	3	US-09-280-799-133	Sequence 133, App
c 38	8	3.5	1001	4	US-09-641-638-113	Sequence 113, App
c 39	8	3.5	1001	4	US-09-641-638-114	Sequence 114, App
c 40	8	3.5	1002	4	US-09-465-353-1	Sequence 1, Appli
c 41	8	3.5	1050	3	US-08-946-026-59	Sequence 59, Appli
c 42	8	3.5	1071	4	US-09-252-991A-9085	Sequence 9085, Ap
c 43	8	3.5	1104	4	US-09-252-991A-8695	Sequence 8695, Ap
c 44	8	3.5	1280	3	US-09-188-930-246	Sequence 246, App
c 45	8	3.5	1280	4	US-09-312-283C-246	Sequence 246, App

ALIGNMENTS

RESULT 1
US-09-315-689-6
; Sequence 6, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Folkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-315-689-6

Alignment Scores:
Pred. No.: 3.16e-16 Length: 534
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 4 Gaps: 0

US-09-938-391-2 (1-230) x US-09-315-689-6 (1-534)

QY	87	GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle	106
DB	109	GGGCTGGCGGGACCTTCGCGCCCTTCCTGCTCGCGCTGCAGGACCTGTACACATC	168
QY	107	ValArgArgAlaAspArg	112
DB	169	GTGCGCGGTGCGCGCGC	186

THIS PAGE BLANK (USPTO)

DR P-PSDB; AAB30493.
XX
XX Preparation of soluble recombinant endostatin involves transforming
PT Streptomyces host with expression vector comprising nucleotide
PT sequence encoding endostatin operably linked to linker and leader
PT peptide -
XX
XX Example 1; Fig 6; 57pp; English.
XX
XX The present sequence encodes human endostatin. The protein is expressed
CC in Streptomyces. Leader sequences of Streptomyces sp. strain C5
CC SnpA and S. venezuelae alpha-amylase proteins are linked to the
CC N-terminal of endostatin. This ensures that endostatin protein is
CC produced as a secreted, soluble protein which needs no refolding, is
CC stable in the fermentation broth and is produced in large quantities.
CC The method is used for preparing soluble recombinant human, murine or
CC primate endostatin, which is useful in the treatment of cancer,
CC inhibition of tumour growth, inhibition of angiogenesis, isolation of
CC receptors for endostatin and for identification of anti-angiogenic
CC compounds in assays. The endostatin protein is produced as a secreted,
CC soluble protein which needs no refolding, is stable in the fermentation
CC broth and is produced in large quantities. Streptomyces are amenable
CC for cultivation in large fermentations allowing for large quantities of
CC soluble endostatin to be produced.
XX
SQ Sequence 552 BP; 83 A; 196 C; 179 G; 94 T; 0 other;

Alignment Scores:
Pred. No.: 5,21e-15 Length: 552
Score: 26400 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 21 Gaps: 0

US-09-938-391-2 (1-230) x AAC62023 (1-552)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIle 106
Db 121 GGCTGGGGGGACCTTCGGCCCTTCCTGTCTCGGCTGCAGACCTGTACACATC 180

QY 107 ValArgAlaAspArg 112
Db 181 GTGGCGCGTGGCGACCGC-198

Search completed: August 18, 2003, 00:50:02
Job time : 274.778 secs

CC the retina, Drusen, endogenous Candida endophthalmitis,
 CC extrapapillary hamartoma of the retinal pigmented epithelium,
 CC fundus flavimaculatus, idiopathic, macular hole, malignant
 CC melanoma, membranoproliferative glomerulonephritis (type II),
 CC metallic intraocular foreign body, morning glory disc syndrome, ora
 CC multiple evanescent white-dot syndrome, neovascularisation of ora
 CC serrata, operating microscope burn, optic nerve head pits,
 CC photocoagulation, punctate inner choroidopathy, rubella,
 CC sarcoidosis, serpiginous or geographic choroiditis, subretinal
 CC fluid drainage, tiled disc syndrome, Toxoplasma retinochoroiditis,
 CC tuberculosis, Vogt-Koyanagi-Harada syndrome, diabetic retinopathy,
 CC non-diabetic retinopathy, brain vein occlusion, central retinal
 CC vein occlusion, retinopathy in premature infants, rubeosis iridis,
 CC neovascular glaucoma, perifoveal telangiectasis, sickle cell
 CC retinopathy, Eale's disease, retinal vasculitis, Von Hippel
 CC Lindau disease, radiation retinopathy, retinal cryoinjury,
 CC retinitis pigmentosa, retinochoroidal coloboma, corneal
 CC neovascularisation due to herpes simplex keratitis, corneal
 CC keratoplasty, pterygia and trauma (all claimed).
 XX
 SQ Sequence 551 BP; 83 A; 195 C; 179 G; 94 T; 0 other;

Alignment Scores:
 Pred. No.: 5.2e-15 Length: 551
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 11.30% Indels: 0
 DB: 24 Gaps: 0

US-09-938-391-2 (1-230) x ABQ81193 (1-551)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
 DB 120 GGGCTGGCGGGACCTTCGGCCCTTCCTGCTCGCGCTGCAGACCTGTACAGCATC 179
 QY 107 ValArgArgAlaAspArg 112
 DB 180 GTGGCGCGTGGCGACCGC 197

RESULT 14
 AAX35375
 ID AAX35375 standard; DNA; 552 BP.
 XX
 AC AAX35375;
 XX
 DT 16-JUL-1999 (first entry)
 XX
 DE SEQ ID 50 of W09916889.
 XX

XX Angiostatin; endostatin; interferon; thrombospondin;
 KW interferon-inducible protein; platelet factor 4; anti-angiogenic;
 KW anti-tumor; multifunctional protein; angiogenic-mediated disease;
 KW cancer; diabetic retinopathy; macular degeneration; arthritis;
 KW tumor cell production; ss.
 XX

OS Homo sapiens.
 XX

XX W09916889-A1.
 XX

XX 08-APR-1999.
 XX

XX 30-SEP-1998; 98WO-US20464.
 XX

XX 01-OCT-1997; 97US-0060609.
 XX

XX (SEAR) SEARLE & CO G D.
 XX

XX Bolanowski MA, Caparon MH, Casperson GF, Gregory SA;
 PI Klein BK, McKearn JP;
 XX
 XX WPI; 1999-255098/21.
 XX

PT New multifunctional proteins useful for treating angiogenic-mediated
 PT diseases
 XX
 PS Disclosure; Page 85; 121pp; English.
 XX
 CC The specification describes multifunctional proteins which comprise
 CC combinations of angiostatin, endostatin, interferon, thrombospondin,
 CC interferon-inducible protein and platelet factor 4, and have
 CC anti-angiogenic and/or anti-tumor activity. The multifunctional protein
 CC may exhibit useful properties such as having similar or greater
 CC biological activity when compared to a single factor or by having
 CC improved half-life or decreased adverse side effects, or a combination
 CC of these properties. The proteins can be used for treating an
 CC angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular
 CC degeneration, or arthritis. They can also be used for inhibiting the
 CC production of tumor cells (characteristic of lung, breast, ovarian,
 CC prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma,
 CC hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor
 CC growth. The present sequence is used in the course of the invention.
 XX
 SQ Sequence 552 BP; 83 A; 196 C; 179 G; 94 T; 0 other;

Alignment Scores:
 Pred. No.: 5.21e-15 Length: 552
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 11.30% Indels: 0
 DB: 20 Gaps: 0

US-09-938-391-2 (1-230) x AAX35375 (1-552)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
 DB 121 GGGCTGGCGGGACCTTCGGCCCTTCCTGCTCGCGCTGCAGACCTGTACAGCATC 180
 QY 107 ValArgArgAlaAspArg 112
 DB 181 GTGGCGCGTGGCGACCGC 198

RESULT 15
 AAC62023
 ID AAC62023 standard; DNA; 552 BP.
 XX
 AC AAC62023;
 XX
 DT 06-MAR-2001 (first entry)
 XX
 DE Nucleotide sequence of human endostatin encoded by plasmid pWALch#15.
 XX

XX Streptomyces sp. strain C5; SnpA; S. venezuelae; alpha-amylase;
 KW endostatin; cancer; tumour growth; angiogenesis; ss.
 XX

OS Homo sapiens.
 XX

XX Key Location/Qualifiers
 XX CDS 1..552
 XX /*tag= a
 XX /product= "endostatin"
 XX

XX W0200060945-A1.
 XX

XX 19-OCT-2000.
 XX

XX 12-APR-2000; 2000WO-US09747.
 XX

XX 13-APR-1999; 99US-0129084.
 XX

XX (MERI) MERCK & CO INC.
 XX

XX Desanti CL, Strohl WR;
 XX
 XX WPI; 2000-686970/67.
 XX

```
US-09-938-391-2 (1-230) x AAS00867 (1-549)
QY      87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIle 106
Db      121 GGCGCTGGCGGGCACCTTCGCGCCTTCCTGCTCGCGCTGCAGACCTGTACAGCATC 180
QY      107 ValArgArgAlaAspArg 112
Db      181 GTGGCGCGTGGCGACCGC 198

RESULT 12
ABA00774
ID ABA00774 standard; cDNA; 549 BP.
XX
AC ABA00774;
XX
DT 18-MAR-2003 (first entry)
XX
DE Human endostatin coding sequence.
XX
KW Gene; human; plasminogen; angiostatin; neovascularisation;
KW kringle domain; cell proliferation; viral vector;
KW replication-defective; cancer; tumour; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..549
FT /*tag= a
FT /product= "Endostatin"
XX WO20028173-A2.
XX 07-NOV-2002.
XX
XX 29-APR-2002; 2002WO-US13461.
XX
XX 30-APR-2001; 2001US-37673P.
XX 05-APR-2002; 2002US-370634P.
XX
XX (CELL-) CELL GENESYS INC.
XX
XX Chang B, Wu WW, Macarthur J, Patel S, Jooss K, Mendez M;
XX WPI; 2003-129131/12.
XX P-PSDB; AAG79753.
XX
XX New recombinant viral vector expressing human angiostatin useful for
XX inhibiting angiogenesis in a mammalian subject with cancer or tumor
XX
XX Example 4; Page 80-82; 83pp; English.
XX
XX This sequence encodes endostatin. Endostatin is a 20 kD C-terminal
XX fragment of collagen XVIII that inhibits angiogenesis. The endostatin
XX coding sequence may be used in the recombinant viral vector of the
XX invention for obtaining angiostatin activity. The vector comprises a
XX promoter capable of expressing human angiostatin operably linked to a
XX structural gene encoding one or more domains of human angiostatin.
XX The vector, which may be a replication-defective viral vector, is useful
XX for inhibiting angiogenesis in a mammal, especially with cancer or a
XX tumour.
XX
XX Note: This sequence is given incorrectly in the sequence listing of
XX the specification as an amino acid sequence.
XX
SQ Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;

Alignment Scores:
Pred. No.: 5,18e-15 Length: 549
Score: 26.00 Matches: 26
Percent Similarity: 100.00%
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
```

```
DB: 25 Gaps: 0
US-09-938-391-2 (1-230) x ABA00774 (1-549)
QY      87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIle 106
Db      121 GGCGCTGGCGGGCACCTTCGCGCCTTCCTGCTCGCGCTGCAGACCTGTACAGCATC 180
QY      107 ValArgArgAlaAspArg 112
Db      181 GTGGCGCGTGGCGACCGC 198

RESULT 13
ABQ81193
ID ABQ81193 standard; cDNA; 551 BP.
XX
AC ABQ81193;
XX
DT 05-DEC-2002 (first entry)
XX
DE Human endostatin coding sequence.
XX
KW Endostatin; human; ophthalmological; ocular neovascularisation;
KW choroidal neovascularisation; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..551
FT /*tag= a
FT /partial
FT /product= "Endostatin"
FT /transl_except= "(pos:2..3,aa:His)"
FT /note= "the CDS does not include a start codon"
XX
XX WO200267971-A2.
XX
XX 06-SEP-2002.
XX
XX 21-FEB-2002; 2002WO-US05336.
XX
XX 22-FEB-2001; 2001US-270787P.
XX 04-APR-2001; 2001US-281296P.
XX
XX (NOVS ) NOVARTIS AG.
XX
XX Brazzell RK, Campochiaro PA, Dixon KH;
XX WPI; 2002-698636/75.
XX P-PSDB; ABB79901.
XX
XX Treating or preventing choroidal neovascularization comprises
XX increasing the amount of endostatin in ocular tissues of afflicted
XX individuals to a choroidal neovascularization inhibiting level -
XX
XX Claim 27; Page 39-40; 44pp; English.
XX
XX The present sequence is a partial coding sequence for human
XX endostatin. A claimed method for the treatment of ocular
XX neovascularisation, especially choroidal neovascularisation,
XX involves increasing the level of endostatin in ocular tissue,
XX especially where the endostatin is encoded by the present sequence,
XX or is its fragment, derivative or variant. The increase is
XX effected by administering a viral vector, especially an adenovirus,
XX adeno-associated virus, a retrovirus or lentivirus vector,
XX comprising an endostatin-encoding nucleic acid. Cells secreting
XX endostatin may be encapsulated and implanted within an individual.
XX The method is used when ocular neovascularisation is caused by
XX histoplasmosis, pathological myopia, angioid streaks, anterior
XX ischaemic optic neuropathy, bacterial endocarditis, Best's disease,
XX birdshot retinopathy, choroidal haemangioma, choroidal
XX naevi, choroidal nonperfusion, choroidal osteomas, choroidal
XX rupture, choroideaemia, chronic retinal detachment, coloboma of
```

CC activity, or combinations of them. The fusion protein (immunofusin) is
 CC used to inhibit angiogenesis and to treat diseases or conditions mediated
 CC by angiogenesis. Conditions that may be treated include solid tumours,
 CC blood born tumours, tumour metastasis, benign tumours including
 CC haemangiomas, acoustic neuromas, neurofibromas, trachomas and pyrogenic
 CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases
 CC e.g. diabetic retinopathy, retinopathy of prematurity, macular
 CC degeneration, corneal graft rejection, neovascular glaucoma, retrolental
 CC fibroplasia, rubeosis and Osler-Weber syndrome; myocardial angiogenesis,
 CC plaque neovascularisation, telangiectasia, haemophilic joints,
 CC angiobroma, wound granulation, and excessive or abnormal stimulation of
 CC endothelial cells, intestinal cells, atherosclerosis, scleroderma and
 CC hypertrophic scars, i.e. keloid scars. The DNA constructs may be used
 CC in gene therapy. The present sequence is a cDNA encoding human
 CC endostatin used in the construction of immunofusin containing human
 CC immunoglobulin gamma (IgG) Fc fragment.
 CC Note: This sequence is stated in claim 12 as being amino acid sequence
 CC of plasminogen fragment, however, the rest of the specification
 CC refers to this sequence as being nucleotide sequence of human endostatin.
 XX
 SQ Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;

Alignment Scores:

Pred. No.: 5,18e-15 Length: 549
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 11.30% Indels: 0
 DB: 21 Gaps: 0

US-09-938-391-2 (1-230) x AA251291 (1-549)

QY 87 GlyLeuAlaClyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
 DB 121 GGCTGGCGGACCTTCGCGGCTTCCTGCTCGCGCTCGAGACCTGTACAGCATC 180
 QY 107 ValArgArgAlaAspArg 112
 DB 181 GTGCGCGTGGCGACCGC 198

RESULT 11

AA500867
 ID AA500867 standard; DNA; 549 BP.

AC AA500867;

DT 04-JUL-2001 (first entry)

XX Human gene fragment encoding Endostatin(TM) protein.

XX Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;
 KW blood borne tumour; leukaemia; tumour metastasis; benign tumour;
 KW haemangioma; acoustic neuroma; neurofibroma; trachoma; rubeosis;
 KW pyrogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
 KW retinopathy of prematurity; macular corneal graft rejection;
 KW neovascular glaucoma; retrolental fibroplasia; Osler-Weber Syndrome;
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
 KW haemophilic joint; angiobroma; wound granulation; ds.

XX Homo sapiens.

PH Key Location/Qualifiers

FT CDS 1..540

FT /product= "Endostatin(TM) C-terminus minus 3"

FT /partial

FT /note= "Variant produced during fermentation
 FT reaction of Pichia pastoris harbouring an expression
 FT plasmid containing the present sequence"

FT CDS 1..543

FT /tag= b

FT /product= "Endostatin(TM) C-terminus minus 2"

FT /partial
 FT /note= "Variant produced during fermentation
 FT reaction of Pichia pastoris harbouring an expression
 FT plasmid containing the present sequence"
 FT CDS 1..546
 FT /tag= c
 FT /product= "Endostatin(TM) C-terminus minus 1"
 FT /partial
 FT /note= "Variant produced during fermentation
 FT reaction of Pichia pastoris harbouring an expression
 FT plasmid containing the present sequence"
 FT CDS 1..549
 FT /tag= d
 FT /product= "Endostatin(TM)"
 FT /partial
 FT /note= "None of the above CDSs have start or stop codons"

WO200119989-A2.

XX 22-MAR-2001.

XX 14-SEP-2000; 2000WO-US25166.

XX 14-SEP-1999; 99US-0153698.

(ENTR-) ENTREMED INC.

Liang H, Sim KL, Chang-Murad A, Zhou X, Madsen J, Boerner RJ;
 Bermejo LL, Mistry FR, Shepard SR, Schrimsher JL;

WPI; 2001-244802/25.

P-PSDB; AAU00896, AAU00897, AAU00898, AAU00899.

Producing Endostatin protein for treating angiogenesis mediated
 diseases such as solid tumours, comprises recombinantly producing the
 protein using an expression system, and recovering and purifying the
 protein

Claim 6; Page 29; 67pp; English.

The sequence encodes Human Endostatin(TM). The new method of the
 invention is useful for producing, recovering and purifying Endostatin
 (TM) from biological sources, such as biological fluids, tissues, cells,
 culture media, and fermentation media. Endostatin(TM) is useful for
 treating angiogenesis mediated diseases such as solid tumours, blood
 borne tumours, leukaemias, tumour metastases, benign tumours, e.g.
 haemangioma, acoustic neuromas, neurofibromas, trachomas, and pyrogenic
 granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases,
 e.g., diabetic retinopathy, retinopathy of prematurity, macular
 degeneration, corneal graft rejection, neovascular glaucoma, colon
 cancer, retrolental fibroplasia, rubeosis, Osler-Weber Syndrome,
 myocardial angiogenesis, plaque neovascularisation, telangiectasia,
 haemophilic joints, angiobroma, and wound granulation. Endostatin(TM)
 is also useful for treating disease of excessive or abnormal stimulation
 of endothelial cells such as intestinal adhesions, atherosclerosis,
 scleroderma and hypertrophic scars. Higher yields of more purified, and
 biologically active Endostatin(TM) are obtained by the new method.
 Endostatin(TM) can be stored in buffers for extended periods of time, and
 also subjected to lyophilisation, while preserving biological activity.
 Centrifugation of broth from fermentation steps in production is avoided,
 preventing unwanted potential cellular lysis and contamination with
 additional proteins, pigments, enzymes and other cellular chemicals and
 debris.

SQ Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;

Alignment Scores:

Pred. No.: 5,18e-15 Length: 549
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 11.30% Indels: 0
 DB: 22 Gaps: 0

CC sufficiently attenuated for use in human gene therapy. The products of
CC the invention have anti-angiogenic, cytostatic, anti-diabetic and
CC ophthalmological activity. The vector is used in gene therapy for
CC inhibiting tumour growth in humans harbouring a solid tumour. The vector
CC expresses an anti-angiogenic polypeptide. An additional use comprises
CC treatment of diabetic retinopathy, where the anti-angiogenic polypeptide
CC inhibits angiogenesis in the vicinity of the retina. The vector is
CC administered to cells ex vivo and then administered to the patient.

XX
XX Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;

Alignment Scores:

Pred. No.: 5,18e-15 Length: 549
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 20 Gaps: 0

US-09-938-391-2 (1-230) x AAX7719 (1-549)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
DB 121 GGCGCTGGCGGGCACCTTCGCGCCCTTCCTGTCTCGCGCTGCAGGACCTGTACAGCATC 180
QY 107 ValArgArgAlaAspArg 112
DB 181 GTGCGCGGTGCGCACCGC 198

RESULT 9

AAA29884
ID AAA29884 standard; cDNA; 549 BP.

XX
XX AAA29884;

XX 22-AUG-2000 (first entry)

DE Human angiogenesis inhibiting factor 1 encoding cDNA.

XX Human; angiogenesis inhibiting factor 1; IAF-1; tumour; antibody;
KW abnormal vessel disease; ss.

XX Homo sapiens.

XX CN1244536-A.

XX 16-FEB-2000.

XX 10-AUG-1998; 98CN-0117150.

XX 10-AUG-1998; 98CN-0117150.

XX (ONCO-) INST ONCOLOGY UNDER TUMOR HOSPITAL CHINE.

XX Yang Z, Guo W;

XX WPI; 2000-388168/34.

XX P-PSDB; AAY90771.

XX Angiogenesis inhibiting factor 1 and its derivative useful for treating
tumors -

XX Example 1; Fig 5; 41pp; Chinese.

XX The present sequence encodes an angiogenesis inhibiting factor (I),
CC designated IAF-1. The present invention also describes: (1) preparation
CC of (I) and its derivative; (2) an IAF binding acceptor and its
CC preparation; and (3) an IAF antibody. (I) is useful for preparing new
CC biological preparations for effectively treating various tumours and
CC abnormal-vessel diseases. The IAF antibody is preferably a polyclonal
CC antibody, mosaic antibody, single stranded antibody and human originated
CC antibody.

SQ Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;

Alignment Scores:

Pred. No.: 5,18e-15 Length: 549
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 21 Gaps: 0

US-09-938-391-2 (1-230) x AAA29884 (1-549)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
DB 121 GGCGCTGGCGGGCACCTTCGCGCCCTTCCTGTCTCGCGCTGCAGGACCTGTACAGCATC 180
QY 107 ValArgArgAlaAspArg 112
DB 181 GTGCGCGGTGCGCACCGC 198

RESULT 10

AAZ51291

ID AAZ51291 standard; cDNA; 549 BP.

XX
XX AAZ51291;

XX 06-JUN-2000 (first entry)

DE Human angiogenesis inhibitor, endostatin cDNA.

XX Human; immunoglobulin gamma Fc fragment; endostatin; immunofusin;
KW angiogenesis; inhibitor; cytostatic; antirheumatoid; antiarthritic;
KW antipsoriatic; antidiabetic; ophthalmological; immunosuppressant;
KW vasotropic; vulnery; treatment; antiarteriosclerosis; tumour;
KW metastasis; atherosclerosis; psoriasis; rheumatoid arthritis;
KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
KW wound granulation; keloid scar; gene therapy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FH CDS 1..549

FT /*tag= a

FT /product= "Endostatin"

FT /note= "Does not include stop codon"

FT /partial

XX WO200011033-A2.

XX 02-MAR-2000.

XX 25-AUG-1999; 99WO-US19329.

XX 25-AUG-1998; 98US-0097883.

XX (LEXI-) LEXINGEN PHARM CORP.

XX Lo K, Li Y, Gillies SD;

XX WPI; 2000-237616/20.

XX P-PSDB; AAY70252.

XX Novel fusion protein of angiostatin or endostatin and an immunoglobulin
PT FC region, useful for treating conditions mediated by angiogenesis,
PT such as rheumatoid arthritis, tumors and macular degeneration -

XX Claim 12; Pages 40-41; 68pp; English.

XX The patent discloses a DNA molecule encoding a fusion protein comprising
CC a signal sequence, an immunoglobulin Fc region, and an angiogenesis
CC inhibitor selected from angiostatin, endostatin, a plasminogen fragment
CC having angiostatin activity, a collagen XVIII fragment having endostatin

SQ Sequence 537 BP; 79 A; 189 C; 176 G; 93 T; 0 other;

Alignment Scores:

Pred. No.:	5,08e-15	Length:	537
Score:	26.00	Matches:	26
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	11.30%	Indels:	0
DB:	22	Gaps:	0

US-09-938-391-2 (1-230) x AAS00868 (1-537)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
 |||||
 Db 109 GGGCTGGCGGCACCTTCGCGCCTTCCTGCTCGCGCTGCAGACCTGTACAGCATC 168

QY 107 ValArgArgAlaAspArg 112
 |||||
 Db 169 GTGCGCGGTGCGGACCGC 186

RESULT 7

AA27004

ID AAA27004 standard; DNA; 546 BP.

AC AAA27004;

XX

DT 11-AUG-2000 (first entry)

XX

DE Human endostatin gene.

XX

KW Human; endothelial cell proliferation inhibitor; collagen XVIII;
 angiogenesis inhibitor; anti-tumour; cytostatic; antiproliferative;
 vasotropic; dermatological; ophthalmological; vulnary;
 antiarteriosclerotic; antidiabetic; haemostatic; contraceptive;
 ocular angiogenic disease; atherosclerosis; scleroderma;
 myocardial angiogenesis; telangiectasia; angiofibroma;
 wound granulation, ds.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..546

FT /*tag= a

FT /product= "Endostatin protein"

XX

PN W0200026368-A2.

XX

PD 11-MAY-2000.

XX

XX

PF 01-NOV-1999; 99WO-US25605.

XX

PR 30-OCT-1998; 98US-0706343.

DR 20-MAY-1999; 99US-0315689.

XX

PA (CHIL-) CHILDRENS MEDICAL CENT.

XX

XX

PI O'Reilly MS, Folkman MJ;

XX

DR WPI; 2000-365617/31.

DR P-PSDB; AAY94323.

XX

XX Novel endostatin capable of inhibiting endothelial cell proliferation
 PT and angiogenesis, useful for treating angiogenesis-dependent cancers
 PT and as birth control agents

XX

PS Claim 10; Page 39; 68pp; English.

XX

XX The present sequence encodes an endostatin protein which is the carboxy
 CC terminal protein of human collagen XVIII. Recombinant mouse endostatin
 CC (20 mg/kg) was administered subcutaneously to mice implanted with Lewis
 CC lung carcinomas. There was tumour mass regression non-detectable levels
 CC after 12 days of therapy due to the angiogenesis inhibitory activity of
 CC endostatin. Thus the protein is useful for treatment of angiogenesis-

CC dependent cancers. The polynucleotide and polypeptide sequences of this
 CC endostatin are useful for treating and diagnosis of tumours, ocular
 CC angiogenic diseases, Osler-Weber syndrome, myocardial angiogenesis,
 CC plaque neovascularisation, telangiectasia, haemophilic joints, and
 CC angiofibroma and wound granulation, for treatment of diseases related to
 CC excessive or abnormal stimulation of endothelial cells e.g. intestinal
 CC adhesions, atherosclerosis, scleroderma. The protein may also be useful
 CC as a birth control agent by reducing or preventing uterine
 CC vascularisation. The gene for endostatin may be isolated from cells or
 CC tissue that express high levels of endostatin, eg. tumour cells, by
 CC generating cDNA from mRNA using reverse transcriptase and then amplifying
 CC the DNA sequence.

XX

SQ Sequence 546 BP; 80 A; 196 C; 177 G; 93 T; 0 other;

Alignment Scores:

Pred. No.:	5.15e-15	Length:	546
Score:	26.00	Matches:	26
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	11.30%	Indels:	0
DB:	21	Gaps:	0

US-09-938-391-2 (1-230) x AAA27004 (1-546)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
 |||||
 Db 121 GGGCTGGCGGCACCTTCGCGCCTTCCTGCTCGCGCTGCAGACCTGTACAGCATC 180

QY 107 ValArgArgAlaAspArg 112
 |||||
 Db 181 GTGCGCGGTGCGGACCGC 198

RESULT 8

AA27719

ID AAX77719 standard; DNA; 549 BP.

XX

AC AAX77719;

XX

DT 10-AUG-1999 (first entry)

XX

DE Human endostatin DNA coding region fragment.

XX

KW Plasminogen; human; angiostatin; endostatin; gene therapy; vector;
 anti-angiogenic; attenuation; cytostatic; anti-diabetic; ophthalmology;
 tumour growth; solid tumour; diabetic retinopathy; retina; ss.

XX

OS Homo sapiens.

XX

PN W09926480-A1.

XX

PD 03-JUN-1999.

XX

PF 20-NOV-1998; 98WO-US24950.

XX

PR 20-NOV-1997; 97US-0975424.

XX

XX (GENE-) GENETIX PHARM INC.

PA (NASI) MASSACHUSETTS INST TECHNOLOGY.

XX

PI Bachelot T, Leboulch P, Pawliuk RJ;

XX

DR WPI; 1999-357696/30.

DR P-PSDB; AAY08693.

XX

XX Anti-angiogenic gene therapy vectors

XX

PS Disclosure; Page 74; 83pp; English.

XX

XX This invention describes a novel viral gene therapy vector comprising a
 CC nucleic acid molecule encoding an anti-angiogenic polypeptide chosen
 CC from human or murine angiostatin, human or murine endostatin and
 CC angiogenesis-inhibiting fusions and fragments, where the viral vector is

PF 01-NOV-1999; 99WO-US25605.
 XX
 PR 30-OCT-1998; 98US-0106343.
 PR 20-MAY-1999; 99US-0315689.
 XX
 PA (CHIL-) CHILDRENS MEDICAL CENT.
 XX
 PI O'Reilly MS, Folkman MJ;
 XX
 DR WPI: 2000-365617/31.
 DR P-PSDB; AAY94324.
 XX
 PT Novel endostatin capable of inhibiting endothelial cell proliferation
 PT and angiogenesis, useful for treating angiogenesis-dependent cancers
 PT and as birth control agents
 XX
 PS Claim 11; Page 40; 68pp; English.
 XX
 CC The present sequence encodes an alternate functional endostatin
 CC protein. When the human endostatin gene sequence AAA27004 is
 CC recombinantly expressed, an observable doublet of protein results, both
 CC versions of which are functional endostatin proteins. The present gene
 CC sequence encodes an endostatin variant which is the same as the protein
 CC encoded by AAA27004 minus the first four amino acids. Recombinant mouse
 CC endostatin (20 mg/kg) was administered subcutaneously to mice implanted
 CC with Lewis lung carcinomas. There was tumour mass regression
 CC non-detectable levels after 12 days of therapy due to the angiogenesis
 CC inhibitory activity of endostatin. Thus the protein is useful for
 CC treatment of angiogenesis-dependent cancers. The polynucleotide and
 CC polypeptide sequences of this endostatin are useful for treating and
 CC diagnosis of tumours, ocular angiogenic diseases, Osler-Webber syndrome,
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,
 CC haemophilic joints, angiofibroma and wound granulation, for treatment of
 CC diseases related to excessive or abnormal stimulation of endothelial
 CC cells e.g. intestinal adhesions, atherosclerosis, scleroderma. The
 CC protein may also be useful as a birth control agent by reducing or
 CC preventing uterine vascularisation. The gene for endostatin may be
 CC isolated from cells or tissue that express high levels of endostatin, eg.
 CC tumour cells, by generating cDNA from mRNA using reverse transcriptase
 CC and then amplifying the DNA sequence.
 XX
 SQ Sequence 534 BP; 77 A; 189 C; 175 G; 93 T; 0 other;

Alignment Scores:
 Pred. No.: 5,05e-15 Length: 534
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 11.30% Indels: 0
 DB: 21 Gaps: 0

US-09-938-391-2 (1-230) x AAA27005 (1-534)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrsIle 106
 |||||
 Db 109 GGGCTGGCGGACCTTCGGCGCTTCTCTCTCTGGCGCTGCGACCTGTACAGCATC 168
 |||||
 QY 107 ValArgArgAlaAspArg 112
 |||||
 Db 169 GTGGCGCGTCCGACCGC 186
 |||||

RESULT 6
 AAS00868
 ID AAS00868 standard; DNA; 537 BP.
 XX
 AC AAS00868;
 XX
 DT 04-JUL-2001 (first entry)
 XX
 DE Human gene fragment encoding Endostatin(TM) N-terminal mutant protein.
 XX
 KW Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;
 KW blood borne tumour; leukaemia; tumour metastasis; benign tumour;

KW haemangioma; acoustic neuroma; neurofibroma; trachoma; rubeosis;
 KW pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
 KW retinopathy of prematurity; macular corneal graft rejection;
 KW neovascular glaucoma; retrolental fibroplasia; Osler-Webber Syndrome;
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
 KW haemophilic joint; angiofibroma; wound granulation; mutant; ds.
 XX
 OS Homo sapiens.
 XX
 XX
 PH Key Location/Qualifiers
 FT CDS 1..534
 FT /tag= a
 FT /product= "Endostatin(TM) N-terminal mutant#1"
 FT /partial
 FT /note= "Variant produced during fermentation
 FT reaction of Pichia pastoris harbouring an expression
 FT plasmid containing the present sequence"
 FT 1..540
 FT CDS
 FT /tag= b
 FT /product= "Endostatin(TM) N-terminal mutant#2"
 FT /partial
 FT /note= "Neither of the above CDSs have start or stop
 FT codons"
 XX
 WO200119989-A2.
 PD 22-MAR-2001.
 XX
 PF 14-SEP-2000; 2000WO-US25166.
 XX
 PR 14-SEP-1999; 99US-0153698.
 XX
 XX (ENTR-) ENTREMED INC.
 PA
 XX
 PI Liang H, Sim KL, Chang-Murad A, Zhou X, Madsen J, Boerner RJ;
 PI Bermejo LL, Mistry FR, Shepard SR, Schrimsher JL;
 XX
 DR WPI: 2001-244802/25.
 DR P-PSDB; AAU00901, AAU00900.
 XX
 PT Producing Endostatin protein for treating angiogenesis mediated
 PT diseases such as solid tumours, comprises recombinantly producing the
 PT protein using an expression system, and recovering and purifying the
 PT protein
 XX
 PS Claim 6; Page 32; 67pp; English.
 XX
 CC The sequence encodes Human Endostatin(TM) N-terminal deletion
 CC mutant lacking the N-terminal 4 amino acids. The new method of the
 CC invention is useful for producing, recovering and purifying Endostatin
 CC (TM) from biological sources, such as biological fluids, tissues, cells,
 CC culture media, and fermentation media. Endostatin(TM) is useful for
 CC treating angiogenesis mediated diseases such as solid tumours, blood
 CC borne tumours, leukaemias, tumour metastases, benign tumours, e.g.
 CC haemangioma, acoustic neuromas, neurofibromas, trachomas, and pyogenic
 CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases,
 CC e.g., diabetic retinopathy, retinopathy of prematurity, macular
 CC degeneration, corneal graft rejection, neovascular glaucoma, colon
 CC cancer, retrolental fibroplasia, rubeosis, Osler-Webber Syndrome,
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,
 CC haemophilic joints, angiofibroma, and wound granulation. Endostatin(TM)
 CC is also useful for treating disease of excessive or abnormal stimulation
 CC of endothelial cells such as intestinal adhesions, atherosclerosis,
 CC scleroderma and hypertrophic scars. Higher yields of more purified, and
 CC biologically active Endostatin(TM) are obtained by the new method.
 CC Endostatin(TM) can be stored in buffers for extended periods of time, and
 CC also subjected to lyophilisation, while preserving biological activity.
 CC Centrifugation of broth from fermentation steps in production is avoided,
 CC preventing unwanted potential cellular lysis and contamination with
 CC additional proteins, pigments, enzymes and other cellular chemicals and
 CC debris.
 XX

PI Sheppard MG, Tong X;
XX WPI: 2002-354068/39.
DR P-PSDB; AA017430.
XX
PT An isolated nucleic acid molecule for the treatment of
PT angiogenesis-related disorder, such as cancers or diabetic retinopathy,
PT encodes an endostatin protein -
XX
PS Claim 2; Fig 4; 56pp; English.
XX
CC The present invention provides the protein and coding sequences of canine
CC pro-endostatin and endostatin. The sequences can be used in the treatment
CC and diagnosis of angiogenesis related disorders, including cancer,
CC rheumatoid arthritis, psoriasis, retinopathy, macular degeneration,
CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasia,
CC rubeosis, Osler-Webber Syndrome, myocardial angiogenesis, plaque
CC neovascularisation, telangiectasia, haemophiliac joints, angiofibroma,
CC wound granulation, coronary collaterals, cerebral collaterals,
CC arteriovenous malformations, ischaemic limb angiogenesis, diabetic
CC neovascularisation, and fractures. The present sequence is the canine
CC endostatin coding sequence.
XX
SQ Sequence 555 BP; 78 A; 203 C; 191 G; 83 T; 0 other;
Alignment Scores:
Pred. No.: 7-7e-163 Length: 555
Score: 184.00 Matches: 184
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.00% Indels: 0
DB: 24 Gaps: 0
US-09-938-391-2 (1-230) x AAL46063 (1-555)
QY 47 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 66
Db 1 CACACCCACCAGGACTTCCAGCTGGTGTGTCACCTGGTGGCCCTGAACAGCCGCGAGCG 60
QY 67 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaA 86
Db 61 GCGCGCATGCGAGGCATCCGGGAGCGGACTTCCAGTGTCTTCAGAGCGCGCGCGCG 120
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 121 GGCTGTGGCGGACCTTCCGGGCTTCTGTGTCTGGCTGCGGCTGACGACCTTACAGCATC 180
QY 107 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 126
Db 181 GTGCGCGCGCGACCGACCGCGGGTGGCGTCTCAACCTCAGGACGAGGTGCTCTTC 240
QY 127 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyAlaArgIle 146
Db 241 CCCAGCTGGAGGCTTATTCGGGCTCCAGGGCAGCTGAAGCCCGGGCGCGCGCATC 300
QY 147 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 166
Db 301 TTCTCTTTCACGCGCAGATGCTTCAGACACCCCGCTGGCCCGGAGACGGTGTGG 360
QY 167 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 186
Db 361 CACGGCTCCGACCCAGCGGCGCGCTGACCGACAGCTACTGCGAGACGTGGCGGACG 420
QY 187 GluAlaProAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 206
Db 421 GAGCGCCCGCGGCCACCGGCGAGGCGTCTGCTGCGCGGCGAGGCTGTGGAGCAG 480
QY 207 GluAlaAspSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 226
Db 481 GAGCGCGGAGCTGCCGCCACGCTTCTGTGTGCTCTGATCATGAGAACGCGCATGACC 540
QY 227 SerPheSerLys 230
|||||

Db 541 TCCTTCTCCAAG 552
RESULT 3
AAZ51309
ID AAZ51309 standard; DNA; 552 BP.
XX
AC AAZ51309;
XX
DT 06-JUN-2000 (first entry)
XX
DE Canine angiogenesis inhibitor, endostatin DNA.
XX
KW Canine; immunoglobulin Fc fragment; endostatin; immunofusin;
KW angiogenesis; inhibitor; cytostatic; antirheumatoid; antiarthritic;
KW antipsoriatic; antidiabetic; ophthalmological; immunosuppressant;
KW vasotropic; vulnary; treatment; antiarteriosclerosis; tumour;
KW metastasis; atherosclerosis; psoriasis; rheumatoid arthritis;
KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
KW wound granulation; keloid scar; gene therapy; ds.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT CDS 1..552
FT /*tag= a
FT /product= "Endostatin"
FT /note= "Does not include stop codon"
FT /partial
XX
PN WO200011033-A2.
XX
PD 02-MAR-2000.
XX
PF 25-AUG-1999; 99WO-US19329.
XX
PR 25-AUG-1998; 98US-0097883.
XX
PA (LEXI-) LEXINGEN PHARM CORP.
XX
PI Lo K, Li Y, Gillies SD;
XX
DR WPI; 2000-237616/20.
DR P-PSDB; AAV70265.
PT Novel fusion protein of angiostatin or endostatin and an immunoglobulin
PT Fc region, useful for treating conditions mediated by angiogenesis,
PT such as rheumatoid arthritis, tumors and macular degeneration -
XX
PS Example 8; Pages 58-59; 68pp; English.
XX
CC The patent discloses a DNA molecule encoding a fusion protein comprising
CC a signal sequence, an immunoglobulin Fc region, and an angiogenesis
CC inhibitor selected from angiostatin, endostatin, a plasminogen fragment
CC having angiostatin activity, a collagen XVIII fragment having endostatin
CC activity, or combinations of them. The fusion protein (immunofusin) is
CC used to inhibit angiogenesis and to treat diseases or conditions mediated
CC by angiogenesis. Conditions that may be treated include solid tumours,
CC blood born tumours, tumour metastasis, benign tumours including
CC haemangiomas, acoustic neuromas, neurofibromas, trachomas and pyrogenic
CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases
CC e.g. diabetic retinopathy, retinopathy of prematurity, macular
CC degeneration, corneal graft rejection, neovascular glaucoma, retrolental
CC fibroplasia, rubeosis and Osler-Webber syndrome; myocardial angiogenesis,
CC plaque neovascularisation, telangiectasia, haemophiliac joints,
CC angiofibroma, wound granulation, and excessive or abnormal stimulation of
CC endothelial cells, intestinal cells, atherosclerosis, scleroderma and
CC hypertrophic scars, i.e. keloid scars. The DNA constructs may be used
CC in gene therapy. The present sequence is a DNA encoding canine
CC endostatin used in the construction of immunofusin containing canine
CC immunoglobulin Fc fragment.
XX
SQ Sequence 552 BP; 77 A; 204 C; 190 G; 81 T; 0 other;

KW ischaemic limb angiogenesis; diabetic neovascularisation; fracture;
 KW cytostatic; antiarthritic; antipsoriatic; antidiabetic; ophthalmological;
 KW gynaecological; gene; ss.
 XX
 OS Canis familiaris.

Key Location/Qualifiers
 CDS 1..693
 FT /*tag= a
 FT /product= "pro-endostatin"
 FT /partial
 FT /note= "no start codon"

EP1191036-A2.

27-MAR-2002.

24-AUG-2001; 2001EP-0307224.

25-AUG-2000; 2000US-227924P.

(PFIZ) PFIZER PROD INC.

Sheppard MG, Tong X;

WPI; 2002-354068/39.

P-PSDB; AAO17429.

XX An isolated nucleic acid molecule for the treatment of
 PT angiogenesis-related disorder, such as cancers or diabetic retinopathy,
 PT encodes an endostatin protein -

XX Claim 2; Fig 2; 56pp; English.

XX The present invention provides the protein and coding sequences of canine
 CC pro-endostatin and endostatin. The sequences can be used in the treatment
 CC and diagnosis of angiogenesis related disorders, including cancer,
 CC rheumatoid arthritis, psoriasis, retinopathy, macular degeneration,
 CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasia,
 CC rubeosis, Osler-Webber Syndrome, myocardial angiogenesis, plaque
 CC neovascularisation, telangiectasia, haemophilic joints, angiobroma,
 CC wound granulation, coronary collateral, cerebral collaterals,
 CC arteriovenous malformations, ischaemic limb angiogenesis, diabetic
 CC neovascularisation, and fractures. The present sequence is the canine
 CC pro-endostatin coding sequence.

XX Sequence 829 BP; 124 A; 314 C; 278 G; 113 T; 0 other;

Alignment Scores:

Pred. No.: 1.02e-205 Length: 829
 Score: 230.00 Matches: 230
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

US-09-938-391-2 (1-230) x AAL46062 (1-829)

QY 1 ProTrpArgAlaAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
 Db 1 CCTGGCGGGCAGATGATCATCTGTGGCGGCCCGCGCTGTGGACCCCGAGCCCTAC 60
 QY 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
 Db 61 CCGGGGGCCCGCACACAGCTCTAGTGCACTTCAGCGGCTCGCCCACTGTGTGG 120
 QY 41 ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuAla 60
 Db 121 CCGGTCCACACCCACACCCACCCAGACTTCAGCTGTGTGACCTGTGTGGCC 180
 QY 61 LeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPhe 80
 Db 181 CTGAACAGCCCGAGCGGGCGCATCGAGGATCCGGGGAGCGGAGCTTCAGTGCTTC 240

QY 81 GlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100
 Db 241 CAGCAGCGCGCGCGCGCGCGCTGGCCGACACTTCCGGGCTTCTGTGCTCGCGGCTG 300
 QY 101 GlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeu 120
 Db 301 CAGGACCTCTACAGCATCTGTGCGCGCGCGCGCGCGCGCGCTGCGTCAACCTC 360
 QY 121 ArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeu 140
 Db 361 AGGACGAGGTGCTCTTCCCGAGCTGGAGGCTTATTCTCGGGCTTCCGAGGCGCAGCTG 420
 QY 141 LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTrp 160
 Db 421 AAGCCCGGGCGCGCATCTTCTTTCGACGCGAGAGATGTCTGCAGACACCCCGCTGG 480
 QY 161 ProArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTyr 180
 Db 481 CCGCGAAGAGCGTGTGGACGCTCCGACCGCGCGCGCGCGCGCGCTGACCGACGCTAC 540
 QY 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla 200
 Db 541 TCGAGACGTGGCGGACGAGCGCGCGCGCGCGCGCGCGCTGCTGTGTGGCG .600
 QY 201 GlyArgLeuLeuGluGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCysIle 220
 Db 601 GGCAGGCTGCTGGAGCAGGAGCGCGCGAGCTGCCGCCACGCCCTTCTGTGTGTCTGTCATC 660
 QY 221 GluAsnSerValMetThrSerPheSerLys 230
 Db 661 GAGAACAGCGTCATGACCTCTCTTCCMAAG 690

RESULT 2

AAL46063

ID AAL46063 standard; cDNA; 555 BP.

XX AAL46063;

DT 19-JUL-2002 (first entry)

XX Canine endostatin coding sequence.

DE Dog; pro-endostatin; endostatin; angiogenesis; cancer; gene therapy;
 KW psoriasis; rheumatoid arthritis; retinopathy; macular degeneration;
 KW corneal graft rejection; neovascular glaucoma; retrolental fibroplasia;
 KW rubeosis; Osler-Webber Syndrome; myocardial angiogenesis;
 KW plaque neovascularisation; telangiectasia; haemophilic joints;
 KW angiofibroma; wound granulation; coronary collateral;
 KW cerebral collateral; arteriovenous malformation;
 KW ischaemic limb angiogenesis; diabetic neovascularisation; fracture;
 KW cytostatic; antiarthritic; antipsoriatic; antidiabetic; ophthalmological;
 KW gynaecological; gene; ss.

OS Canis familiaris.

Key Location/Qualifiers
 CDS 1..555
 FT /*tag= a
 FT /product= "endostatin"
 FT /partial
 FT /note= "no start codon"

EP1191036-A2.

27-MAR-2002.

24-AUG-2001; 2001EP-0307224.

25-AUG-2000; 2000US-227924P.

(PFIZ) PFIZER PROD INC.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 17, 2003, 22:57:56 ; Search time 272.778 Seconds
(without alignments)
2276.105 Million cell updates/sec

Title: US-09-938-391-2

Perfect score: 230

Sequence: 1 PWRADILAGPRLPDQPY.....CRHAFVVLCTIENSVMTSFSK 230

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Word size: 1

Total number of hits satisfying chosen parameters: 5103490

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -DEV=slp
-MODEL=frame_p2n.model -DB=us-09-938-391-2.p2noli.rng -Q=/cg2_1/USP01.spool_p/US09938391/runat_04082003_130738_27574/app_query.fasta_1.718
-DB=N Geneseq 19Jun03 -QFMT=fastap -SUFFIX=p2noli.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938391_@CGN_1_1_605_@runat_04082003_130738_27574 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N Geneseq 19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	230	100.0	829	24	AAL46062	Canine pro-endosta
2	184	80.0	555	24	AAL46063	Canine endostatin
3	176	76.5	552	21	AAZ51309	Canine angiogenesi
4	27	11.7	555	22	AAC88291	Chicken endostatin
5	26	11.3	534	21	AAA27005	Altermate human en
6	26	11.3	537	22	AA500868	Human gene fragmen
7	26	11.3	546	21	AA27004	Human endostatin g
8	26	11.3	549	20	AA77719	Human endostatin D
9	26	11.3	549	21	AA29884	Human angiogenesi
10	26	11.3	549	21	AAZ51291	Human angiogenesi
11	26	11.3	549	22	AA500867	Human gene fragmen
12	26	11.3	549	25	ABA00774	Human endostatin c
13	26	11.3	551	24	ABQ81193	Human endostatin c
14	26	11.3	552	20	AA353375	SEQ ID 50 of WO991
15	26	11.3	552	21	AAC62023	Nucleotide sequenc
16	26	11.3	552	21	AAA68203	Human endostatin c
17	26	11.3	552	21	AAZ50398	Human endostatin c
18	26	11.3	552	22	AAC88289	Human endostatin c
19	26	11.3	552	24	ABK50685	cDNA encoding huma
20	26	11.3	555	22	AA166529	Human vascular end
21	26	11.3	558	24	ABA99261	Human endostatin c
22	26	11.3	563	20	AAZ08750	Human endostatin c
23	26	11.3	574	24	ABV94691	Human pancreatic c
24	26	11.3	641	22	AAH79104	Human endostatin e
25	26	11.3	641	24	AA155454	Specific tumour ce
26	26	11.3	641	24	AA144000	Reproductive recom
27	26	11.3	641	24	ABQ76079	Anticancer gene-as
28	26	11.3	650	24	ABK09978	Synthetic plasmid
29	26	11.3	816	21	AAA64013	DNA encoding a hum
30	26	11.3	900	24	ABK09977	Synthetic plasmid
31	26	11.3	968	21	AAC62025	Nucleotide sequenc
32	26	11.3	1564	24	ABQ76740	DNA encoding human
33	26	11.3	3394	18	AA184484	Human alpha-1 coll
34	26	11.3	3394	20	AA178379	Human alpha (XVII
35	26	11.3	3394	24	ABN95680	Gene #2178 used to
36	26	11.3	4551	24	ABV94763	Human pancreatic c
37	26	11.3	4551	24	ABN85301	Human collagen XVI
38	26	11.3	4875	24	ABQ54955	Human ovarian anti
39	26	11.3	5408	20	AA177720	Human collagen 18
40	26	11.3	6462	21	AAA64014	Nucleotide sequenc
41	26	11.3	37664	22	AAK83781	Human immune/haema
42	26	11.3	61710	22	AAK83782	Human immune/haema
43	23	10.0	552	21	AA177715	Murine endostatin
44	23	10.0	552	21	AAZ51299	Murine angiogenesi
45	23	10.0	555	20	AA184635	Mouse endostatin c

ALIGNMENTS

RESULT 1
AAL46062
ID AAL46062 standard; cDNA; 829 BP.
XX
AC AAL46062;
XX
DT 19-JUL-2002 (first entry)

Canine pro-endostatin coding sequence.

- XX Dog; pro-endostatin; endostatin; angiogenesis; cancer; gene therapy;
- XX psoriasis; rheumatoid arthritis; retinopathy; macular degeneration;
- XX corneal graft rejection; neovascular glaucoma; retrolental fibroplasia;
- XX rubrosis; Osler-Webber Syndrome; myocardial angiogenesis;
- XX plaque neovascularisation; telangiectasia; haemophilic joints;
- XX angiofibroma; wound granulation; coronary collateral;
- XX cerebral collateral; arteriovenous malformation;

THIS PAGE BLANK (USPTO)

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 786)
AUTHORS Dou,D.
TITLE Direct Submission
JOURNAL Submitted (27-JUN-2000) R & D, Robinson Biotech Corp., 24634 Five
Mile Road, Suite 24-26, Redford, MI 48239, USA
FEATURES
source Location/Qualifiers
1..786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
CDS
1..786
/note="antiangiogenic agent"
/codon_start=1
/product="multi-functional protein MFP"
/protein_id="AA052249.1"
/db_xref="GI:21425751"
/translation="MSEGNSSGGSGGGVITKIGGSGGGSGGGRNPDGAKPWGGG
SGGSGGGTTPRKLYDYGGSGGGSGGGSGGGSHRDFQPVHLVALNSPLSGMR
GIRGADFCQARAGLAGTFRFLSLQDLISIVRADRAAIVPLNKDELIFPS
WEALFSGSGPLKPGARIFFDQKQVLRHTPWQKSVHMGSDPNCRRLTESYCETWRT
EAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTASK"
BASE COUNT 124 a 239 c 280 g 143 t
ORIGIN
Alignment Scores:
Pred. No.: 3.39e-17 Length: 786
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 9 Gaps: 0
US-09-938-391-2 (1-230) x AF282883 (1-786)
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 355 GGCGCTGGCGGGACCTTCGCGCCTTCCTGCTCGCGCCTCGAGACCTGTACAGCATC 414
QY 107 ValArgArgAlaAspArg 112
Db 415 GTGCGCGGTGCGGACCGC 432
RESULT 15
AX370851
LOCUS AX370851 900 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 12 from Patent WO0210372.
ACCESSION AX370851
VERSION AX370851.1 GI:19168980
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Chapman,P.W., de Luca,G. and Falciola,L.
TITLE Method of producing functional protein domains
JOURNAL Patent: WO 0210372-A 12 07-FEB-2002;
Applied Research Systems ARS Holding S.A. (AN)
FEATURES
source Location/Qualifiers
1..900
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="mRNA for mIGSP fused to COL18A1 exons 38-41 for
pEnd-HR#1"
BASE COUNT 146 a 336 c 274 g 144 t
ORIGIN
Alignment Scores:
Pred. No.: 3.8e-17 Length: 900

Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 6 Gaps: 0
US-09-938-391-2 (1-230) x AX370851 (1-900)
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 457 GGCGTGGCGGGACCTTCGCGCCTTCCTGCTCGCGCCTCGAGACCTGTACAGCATC 516
QY 107 ValArgArgAlaAspArg 112
Db 517 GTGCGCGGTGCGGACCGC 534
Search completed: August 18, 2003, 02:40:04
Job time : 3665.11 secs


```

/product="type XVIII collagen"
/protein_id="AAF01310.1"
/db_xref="GI:6013265"
/translation="MHSRDFQPVHLVALNSPLSGCMRGIRGADFQCFQARAVGLA
SFGKFLSLRLQDLYSIVRRADAAVPIVNLKDELLFPSEALFSGSEGLPKGARIF
SFGKDLVLTPTWPKQSVHGGDPNGRRLTESYCTWRTEAPSATQAYSLIGGRLLG
QSAASCHHAYIVLICIENSFMTASK"
BASE COUNT      87 a  196 c  177 g  95 t
ORIGIN

Alignment Scores:
Pred. No.:      2,52e-17      Length:      555
Score:          26.00         Matches:      26
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:     11.30%      Indels:          0
DB:              9           Gaps:            0

US-09-938-391-2 (1-230) x AF184060 (1-555)

QY      87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
      |||
Db      124 GGGCTGGCGGACCTTCGGCGCTTCCTGCTCGCGCTGCGACCTGTACAGCATC 183

QY      107 ValArgAlaAspArg 112
      |||
Db      184 GTGGCGCGTGGCGACCGC 201

RESULT 12
AF416592
LOCUS      AF416592
DEFINITION Homo sapiens collagen XVIII mRNA, partial cds.
ACCESSION  AF416592
VERSION     AF416592.1 GI:17226297
KEYWORDS   Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 564)
AUTHORS     Peng,Y., Cui,L.B., Liu,C.X. and Ma,Q.J.
TITLE       Inhibition effect in vitro of purified endostatin expressed in
JOURNAL     Pichia pastoris
MEDLINE     21409408
PUBMED      11517600
REFERENCE   2 (bases 1 to 564)
AUTHORS     Peng,Y., Wu,Y., Zhu,X., Liu,C.X. and Ma,Q.J.
TITLE       Endostatin contributes to maintain cartilage homeostasis via
JOURNAL     promotion of the anabolic program of chondrocytes
REFERENCE   3 (bases 1 to 564)
AUTHORS     Peng,Y., Cui,L.B. and Ma,Q.J.
TITLE       Direct Submission
JOURNAL     Submitted (03-SEP-2001) Genetic Engineering, Beijing Institute of
Biotechnology, Taiping Road, Beijing 100850, P.R. China
FEATURES   source
            1..564
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /chromosome="21"
            <1..564
            /note="similar to endostatin; can inhibit endothelial cell
            proliferation"
            /codon_start=1
            /product="collagen XVIII"
            /protein_id="AAL37720.1"
            /db_xref="GI:17226298"
            /translation="SPPAHSHRDFQPVHLVALNSPLSGCMRGIRGADFQCFQARAV
            GLAGTFAFLSSRLQDLYSIVRRADAAVPIVNLKDELLFPSEALFSGSEGLPKPGA
            RIFSGDKVLRLPTWPKQSVHGGDPNGRRLTESYCTWRTEAPSATQASLLGGR
            LLGQSAASCHHAYIVLICIENSFMTASK"

CDS

```

```

misc_feature <1..561
            /note="Region: NCI domain"
BASE COUNT      86 a  203 c  180 g  95 t
ORIGIN

Alignment Scores:
Pred. No.:      2,55e-17      Length:      564
Score:          26.00         Matches:      26
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:     11.30%      Indels:          0
DB:              9           Gaps:            0

US-09-938-391-2 (1-230) x AF416592 (1-564)

QY      87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
      |||
Db      133 GGGCTGGCGGACCTTCGGCGCTTCCTGCTCGCGCTGCGACCTGTACAGCATC 192

QY      107 ValArgAlaAspArg 112
      |||
Db      193 GTGGCGCGTGGCGACCGC 210

RESULT 13
AX370853
LOCUS      AX370853
DEFINITION Sequence 14 from Patent WO0210372.
ACCESSION  AX370853
VERSION     AX370853.1 GI:19168981
KEYWORDS   synthetic construct
SOURCE      synthetic construct
ORGANISM    artificial sequences.
REFERENCE   1
AUTHORS     Chapman,P.W., de Luca,G. and Falcicola,L.
TITLE       Method of producing functional protein domains
JOURNAL     Patent: WO 0210372-A 14 07-FEB-2002;
            Applied Research Systems ARS Holding S.A. (AN)
FEATURES   Location/Qualifiers
            source
            1..650
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
            /note="mRNA for mlgSP fused to COL18A1 exons 39-41 for
            pEnd-HR#2"
BASE COUNT      104 a  218 c  207 g  121 t
ORIGIN

Alignment Scores:
Pred. No.:      2,88e-17      Length:      650
Score:          26.00         Matches:      26
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:     11.30%      Indels:          0
DB:              6           Gaps:            0

US-09-938-391-2 (1-230) x AX370853 (1-650)

QY      87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
      |||
Db      211 GGGCTGGCGGACCTTCGGCGCTTCCTGCTCGCGCTGCGACCTGTACAGCATC 270

QY      107 ValArgAlaAspArg 112
      |||
Db      271 GTGGCGCGTGGCGACCGC 288

RESULT 14
AF282883
LOCUS      AF282883
DEFINITION Homo sapiens multi-functional protein MFP mRNA, complete cds.
ACCESSION  AF282883
VERSION     AF282883.1 GI:21425750
KEYWORDS

```

```

DB: 6 Gaps: 0
US-09-938-391-2 (1-230) x AX395662 (1-552)
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
DB 121 GGGCTGGCGGGACCTTCGCGCCTTCCTGCTCGCGCTGCAGGACCTGTACAGCATC 180
QY 107 ValArgArgAlaAspArg 112
DB 181 GTGGCGCGTGGCGACCGC 198

RESULT 9
AX473835
LOCUS AX473835 552 bp DNA linear PAT 09-AUG-2002
DEFINITION Sequence 2 from Patent WO0230982.
ACCESSION AX473835
VERSION AX473835.1 GI:22208005
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Sim, K.L. and Liang, H.
TITLE Angiogenesis-inhibiting peptides and proteins and methods of use
JOURNAL Patent: WO 0230982-A 2 18-APR-2002;
EntreMed, Inc.; (US)
FEATURES
source
Location/Qualifiers
1..552
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 83 a 196 c 179 g 94 t
ORIGIN

Alignment Scores:
Pred. No.: 2,51e-17 Length: 552
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 6 Gaps: 0

US-09-938-391-2 (1-230) x AX473835 (1-552)
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
DB 121 GGGCTGGCGGGACCTTCGCGCCTTCCTGCTCGCGCTGCAGGACCTGTACAGCATC 180
QY 107 ValArgArgAlaAspArg 112
DB 181 GTGGCGCGTGGCGACCGC 198

RESULT 10
BD081407
LOCUS BD081407 552 bp DNA linear PAT 27-AUG-2002
DEFINITION Fused protein containing angiotensin component and utilization
thereof in antitumor therapy.
ACCESSION BD081407
VERSION BD081407.1 GI:22627010
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Bolanowski, M.A., Caparon, M.H., Casperson, G.F., Gregory, S.A.,
Klein, B.K. and McKearn, J.P.
TITLE Fused protein containing angiotensin component and utilization
thereof in antitumor therapy
JOURNAL Patent: JP 2001518304-A 50 16-OCT-2001;
GD SEARLE AND CO

```

```

COMMENT OS Homo sapiens, (human)
PN JP 2001518304-A/50
PD 16-OCT-2001
PP 30-SEP-1998 JP 20005113958
PR 01-OCT-1997 US 60/060609
PI MARK A BOLANOWSKI, MAIRE H CAPARON, GERALD F CASPERSON, SUSAN A
PI GREGORY,
PI BARBARA K KLEIN, JOHN P MCKEARN
PC C12N15/09, A61K38/00, A61K48/00, A61P9/10, A61P35/00, C07K14/52, PC
C07K14/56,
PC C07K14/78, C12N9/68, C12N15/00, A61K37/02
CC Fused protein containing angiotensin component and utilization
thereof in
CC antitumor therapy
FH Key Location/Qualifiers
FT source 1..552
/organism="Homo sapiens (human)"
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 83 a 196 c 179 g 94 t
ORIGIN

Alignment Scores:
Pred. No.: 2,51e-17 Length: 552
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.30% Indels: 0
DB: 6 Gaps: 0

US-09-938-391-2 (1-230) x BD081407 (1-552)
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
DB 121 GGGCTGGCGGGACCTTCGCGCCTTCCTGCTCGCGCTGCAGGACCTGTACAGCATC 180
QY 107 ValArgArgAlaAspArg 112
DB 181 GTGGCGCGTGGCGACCGC 198

RESULT 11
AF184060
LOCUS AF184060 555 bp mRNA linear PRI 06-OCT-1999
DEFINITION Homo sapiens type XVIII collagen mRNA, complete cds.
ACCESSION AF184060
VERSION AF184060.1 GI:6013264
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Zhi-Yong, H., Biao, L., Wei-Jie, Z. and Xiang-Fu, W.
TITLE Cloning and expression of human endostatin gene in Escherichia coli
JOURNAL Unpublished
REFERENCE 2
AUTHORS Zhi-Yong, H., Biao, L., Wei-Jie, Z. and Xiang-Fu, W.
TITLE Direct Submision
JOURNAL Submitted (07-SEP-1999) Shanghai Institute of Biochemistry, Chinese
Academy of Sciences, 320 Yueyang Road, Shanghai 200031, P. R. China
FEATURES
source
Location/Qualifiers
1..555
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="placenta"
1..555
/feature="endostatin"
/feature_start=1
CDS
/codon_start=1

```



```
Db 421 GAGGCCCGCGCCACCGGCGAGGCGTCTGCTCGCTCGCGCGCAGGCTGCTGGAGCAG 480
|||||
QY 207 GUUAlAAlaserCysArghisAlaPheValValValLeuCysIleGluAsnSerValMetThr 226
|||||
Db 481 GAGGCCCGCGAGCTGCCCGCACGCTTCTGCTGCTCTGTCATCGAGAACAGCGTCATGACC 540
|||||
QY 227 SerPheSerIys 230
|||||
Db 541 TCCTTCTCAAG 552
|||||

RESULT 3
AF083440
LOCUS      5279 bp      mRNA      linear      VRT 31-AUG-2001
DEFINITION      Gallus gallus collagen XVIII mRNA, complete cds.
ACCESSION      AF083440
VERSION      AF083440.2 GI:15383904
KEYWORDS
SOURCE      Gallus gallus (chicken)
ORGANISM      Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 5279)
Collagen XVIII is a basement membrane heparan sulfate proteoglycan
J. Biol. Chem. 273 (39), 25404-25412 (1998)
98411346
PUBMED 9738008
2 (bases 1 to 5279)
Halfter,W. and Dong,S.
Direct Submission
Submitted (31-AUG-1998) Neurobiology, University of Pittsburgh,
3500 Terrace Street, Pittsburgh, PA 15261, USA
3 (bases 1 to 5279)
Halfter,W. and Dong,S.
Direct Submission
Submitted (31-AUG-2001) Neurobiology, University of Pittsburgh,
3500 Terrace Street, Pittsburgh, PA 15261, USA
JOURNAL
REMARK      Sequence update by submitter
COMMENT      On Aug 31, 2001 this sequence version replaced gi:3493142.
FEATURES
source
1..5279
Location/Qualifiers
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
8..4042
/codon_start=1
/product="collagen XVIII"
/protein_id="AAC33294.2"
/db_xref="GI:15383905"
/translation="MRPGCAPRRLLLLGLFVLLLLPAASQEBENLSTEVSLLELIGDP
PPEELIKYGPNNPGYVSGANTQVARYHLPGFFYRDFSLLFHQPTPRAGMLF
AITDTSQIYYVGLKSLQMGKQIIFYTEPGSSYAAATFTVTLNQTWTFPAI
SVEDRVILYDCEHERVRFERSDEMELRSGSLFVAAGADFPDKVQGYADLR
RGDPAAEQCEEREDDAASGDFSGAEDRHPSGDKXGIPCLLDVAVPTSPVVEG
SSTRSAGSPQAAERTRVERLQVSTGCTGPKGKGEKGERGPKGDSGSLTGAT
KGEKGSLGELGINGSAGFGPGSKGQKGPFGPPGPPPLSRHTDSLSLQVTPGPGT
GPPGKADGEPDGPDEDKPGEMGPGQFPPTGSGQKGEKGDGPGVGRPPGP
PGPPGPRSSQKGLTIFDMGSGFGGDLTLRGRPPGPPGPPGPPGLPGEFRF
GNNSTDLPGPGLPGRDGTPGAPGPEGLGPGKPGPPGPKGERGVGLGLPGA
PGPKSGKAGPAGPGETGLAGLPVGPGRPPGPPGPPGPGCYEAGFGDMGSG
LPLATGSPGPRGPGVGLPGIKGVEGSLGQPGPPGPKGDAGVPGVGRPLEGF
PGQPGKNGRSGPGERGQDGLPFPFPFGPPGPPGYIMSSDRPLVALPGEGR
SCHAGFPVGPKGDPGPGIQCAPGMKEGEPGVIISPDGTIVAANYKQKGFGL
PGMGPSPHGRAGMKETIGFPGRPGMGLKGEKGDVDSISVLSRGPFGPGL
PGPPGPGVSVNNGFSDRAPPAPFPQPGKGEKGDVGAPOQFPYDLRSF
SASLRGDKGEAGPKGKGSSTLGVSTGPPGPGVPGPKGDSIVGPPGPG
PGPPGPGVYEGCGGPPGPPGPPSPFPHPQAISNPGPPGPPGPPGPGTSLG
LRIMPTYQWLGAHLEPGGLIFLADROLRYIRLGRFRRLSEHLTVPSALDNE
VYDKLPSTHYGAQQVHPRLNRNPPTPRMRGDEVANQHHLPQPLLOHLLNS
YTHRPDPAPAAHVHDFQFALHLVALNTLSGMRGIRGADFQCFOQARQVGLAG
TFRAFLSLRLQDLYSIVRRADRTAVPIVNLREVLVFNWEALFTGSEAPLRAGARILS
```

AX399629
LOCUS AX399629 829 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 1 from Patent EP1191036.
ACCESSION AX399629
VERSION AX399629.1 GI:21335409
KEYWORDS
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1
REFERENCE Sheppard, M.G. and Tong, X.
AUTHORS Methods and compositions for diagnosing and treating disorders
TITLE involving angiogenesis
JOURNAL Patent: EP 1191036-A 1 27-MAR-2002;
Pfiizer Products Inc. (US)
FEATURES Location/Qualifiers
source 1..829
/organism="Canis familiaris"
/mol_type="genomic DNA"
/db_xref="taxon:9615"
BASE COUNT 124 a 314 c 278 g 113 t
ORIGIN
Alignment Scores:
Pred. No.: 2.72e-230 Length: 829
Score: 230.00 Matches: 230
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-938-391-2 (1-230) x AX399629 (1-829)
Qy 1 ProTrpArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
Db 1 CCCTGGGGGGGAGATGACATCTTGGCGGGCCCCCGGCGCTGTGGACCCCCCAGCCCTAC 60
Qy 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaAlaArgProThrGlyGly 40
Db 61 CCGGGGCCCCGACACCAGGCTCTTACGTGTGACTTCCAGCCGGCTCGCCCCACTGGTGGG 120
Qy 41 ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuValAla 60
Db 121 CCGCTGCACACCCACACCCACACCCAGGACTTCCAGCTGGTGTCTGCACCTGGTGCC 180
Qy 61 LeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPhe 80
Db 181 CTGAACAGCCCGACGCGGGGGGATGCGAGGCATCCGGGGAGCGGACTTCAGTGTCTTC 240
Qy 81 GlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100
Db 241 CAGCAGGCGCGCGCGGGGCTGGCGGCACTTCCGGGCTTCTCTGTCTGTCTGGGGCTG 300
Qy 101 GlnAspLeuTyrSerIleValArgAlaAlaAspArgThrGlyValProValValAsnLeu 120
Db 301 CAGGACCTCTACAGCATCGTGGCGCGCCGACCCAGCCACCGGGGTGCCGTCTCAACCTC 360
Qy 121 ArgAspGluValValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeu 140
Db 361 AGGGACAGAGGTGCTCTTCCCGAGCTGGGAGGCTTTATTTCTCGGGCTCCGAGGGCCAG 420
Qy 141 LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTrp 160
Db 421 AAGCCCGGGGGCGGCACTTCTCTTTTCAGCGCAGATGTCCTGCAGACACCCCGCTGG 480
Qy 161 ProArgTysSerValTrpHisIleGlySerAspProSerGlyArgArgLeuThrAspSerTyr 180
Db 481 CCCCAGGAAGCGTGTGCACGGCTCCGACCCACCGCGGCGCGCTGTACCCAGCAGCTAC 540
Qy 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla 200
Db 541 TGCAGAGCTGGCGACGAGGCCCGCGGGCCACCGGGCAGGCGTCTGTCTGTCTGGCGG 600

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 17, 2003, 23:00:16 ; Search time 3661.11 Seconds
(without alignments)
2570.043 Million cell updates/sec

Title: US-09-938-391-2

Perfect score: 230

Sequence: 1 PWRADDILAGPRLDPPQY.....CRHAFVLCIENSVMTSFSK 230

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0

Ygapop 60.0 , Ygapext 60.0

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Word size: 1

Total number of hits satisfying chosen parameters: 5773148

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+pn.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0938391/runat_04082003_130738_27585/app_query.fasta_1.718
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2noli.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0938391@cgn_1_15066@runat_04082003_130738_27585 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl: *

1: gb_ba: *

2: gb_htg: *

3: gb_in: *

4: gb_om: *

5: gb_ov: *

6: gb_pat: *

7: gb_ph: *

8: gb_pl: *

9: gb_pr: *

10: gb_ro: *

11: gb_sts: *

12: gb_sy: *

13: gb_un: *

14: gb_vi: *

15: em_ba: *

16: em_fun: *

17: em_hum: *

18: em_in: *

19: em_mu: *

20: em_om: *

21: em_or: *

22: em_ov: *

23: em_pat: *

24: em_ph: *

25: em_pl: *

26: em_ro: *

27: em_sts: *

28: em_un: *

29: em_vi: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rod: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	230	100.0	829	6	AX399629 Sequence
2	184	80.0	555	6	AX399631 Sequence
3	27	11.7	5279	5	AF083440 Gallus ga
4	26	11.3	534	6	AR193166 Sequence
5	26	11.3	537	6	AX100092 Sequence
6	26	11.3	546	6	AR193165 Sequence
7	26	11.3	549	6	AX100086 Sequence
8	26	11.3	552	6	AX395662 Sequence
9	26	11.3	552	6	AX473835 Sequence
10	26	11.3	552	6	BD081407 Fused pro
11	26	11.3	555	9	AF184060 Homo sapi
12	26	11.3	564	9	AF416592 Homo sapi
13	26	11.3	650	6	AX370853 Sequence
14	26	11.3	786	9	AF282883 Homo sapi
15	26	11.3	900	6	AX370851 Sequence
16	26	11.3	2364	9	AX098216 Homo sapi
17	26	11.3	2637	9	BC033715 Homo sapi
18	26	11.3	3394	6	AX409531 Sequence
19	26	11.3	3394	6	IS1045 Sequence 4
20	26	11.3	3394	9	HUMCOL18AX
21	26	11.3	3400	9	AF018082 Homo sapi
22	26	11.3	5929	9	AF018081 Homo sapi
23	26	11.3	131563	9	HSE12310
24	26	11.3	223201	9	HSE3110
25	26	11.3	340000	9	HS21C102
26	23	10.0	515	10	RNO236873
27	23	10.0	552	10	AF257775 Mus muscu
28	23	10.0	555	6	BD102800 Process f
29	23	10.0	555	6	BD128313 Endostati
30	23	10.0	558	6	AR268609 Sequence
31	23	10.0	565	6	AR100642 Sequence
32	23	10.0	565	6	E34073 Carrier/DNA
33	23	10.0	573	6	AR183471 Sequence
34	23	10.0	573	6	AR183520 Sequence
35	23	10.0	573	6	AR217279 Sequence
36	23	10.0	573	6	AR282751 Sequence
37	23	10.0	573	6	AX042272 Sequence
38	23	10.0	573	6	AX490666 Sequence
39	23	10.0	624	6	AX084406 Sequence
40	23	10.0	624	6	AX128409 Sequence
41	23	10.0	639	11	BV071218 S212P6038
42	23	10.0	851	10	AF189709 Rattus no
43	23	10.0	1615	10	MMU03714 Mus musculu
44	23	10.0	3129	10	MMCOL18A11
45	23	10.0	4031	6	IS1044 Sequence 1

ALIGNMENTS

RESULT 1

ORIGIN

Alignment Scores:

Pred. No.:	5,63e-61	Length:	551
Score:	822.00	Matches:	157
Percent Similarity:	94.32%	Conservative:	9
Best Local Similarity:	89.20%	Mismatches:	10
Query Match:	66.34%	Indels:	0
DB:	10	Gaps:	0

US-09-938-391-2 (1-230) x BF074459 (1-551)

```

QY 55 ValLeuHisLeuValAlaLeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGly 74
Db 1 GTGCTGCACCTGGTGGCGCTCAACAGCCGCGAGTCGGGCGGCTCGCGGCGATCCGGCGC 60
QY 75 AlaAspPheGlnCysPheGlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAla 94
Db 61 GCCGACTTTTCAGTGTCTCCAGCAGGCGCGCGCGGGGCTGGCGGCGACCTTCGGCGCG 120
QY 95 PheLeuSerSerArgLeuGlnAspLeuTyrSerIleValArgAlaAspArgThrGly 114
Db 121 TTCCTGTCTCGCGGTTGAGGACCTGTACAGCATCGTGGCGCGCGCCGACCGTCCACC 180
QY 115 ValProValValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSer 134
Db 181 CTGCCCGTCTGTCNACCTCAGGGACGAGGTGCTGTTCTCTAGCTGGGAGGCTTGTCTCA 240
QY 135 GlySerGluGlyGlnLeuLysProGlyAlaArgIlePheSerPheAspGlyArgAspVal 154
Db 241 GGCTCCGAGGGCCAGCTGAAGCCGCGCGCGCGCATCTTCTCTTCGACGCGCAGATGTC 300
QY 155 LeuGlnHisProAlaTrpProArgLysSerValTrpHisGlySerAspProSerGlyArg 174
Db 301 CTTACGATCCCACTGGGCCCCAGAGAGCGGTGTGGCAGCGGCTCATACCCACGCGGCGC 360
QY 175 ArgLeuThrAspSerTyrCysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGln 194
Db 361 CGGCTGACCCGAGAGCTACTCGGAGACGTGGCGGACGACAGCGGCGGCCACAGGCCAG 420
QY 195 AlaSerSerLeuLeuAlaGlyArgLeuLeuGlnGlnAlaAlaSerCysArgHisAla 214
Db 421 GCCTCCTCGTGTGGCGGCGCGGCTGCTGGAGCAGAAAGCCGCGGTGCCACACGCC 480
QY 215 PheValValLeuCysIleGluAsnSerValMetThrSerPheSerLys 230
Db 481 TTCATTGTCTCTGCATCGAGAACAGCTTCATGACCTCTCTCTCCCAAG 528

```

Search completed: August 17, 2003, 22:57:44
 Job time : 2043.33 secs

```

LOCUS      CB596713              715 bp      mRNA      linear      EST 03-APR-2003
DEFINITION AGENCOURT 12990486 NIH MGC 178 Mus musculus cDNA clone
IMAGE:30299651 5', mRNA sequence.
ACCESSION  CB596713
VERSION     CB596713.1  GI:29514569
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 715)
REFERENCE   NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS    National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE       Unpublished
JOURNAL
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgaabs@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
            CDNA Library Preparation: Michael Brownstein Laboratory
            DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
            Clone distribution: Agencourt Bioscience Corporation
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDCM64 row: b column: 12
            High quality sequence stop: 567.
FEATURES    Location/Qualifiers
            1..715
                /organism="Mus musculus"
                /mol_type="mRNA"
                /db_xref="taxon:10090"
                /clone="IMAGE:30299651"
                /lab_host="DH10B (r1-phage-resistant)"
                /note="Organ: lung and heart; Vector: pDNR-LIB; Site 1:
                SfiI (ggccattagcgc); Site 2: SfiI (ggcgctcgcc); cDNA
                made by oligo-dT priming and directionally cloned. 5' and
                3' adaptors were used in cloning as follows:
                5'-AGCAGCTGGTATCAAGCAGAGTGGCCATTACGCCGGG-3' and
                5'-ATTCTAGAGCGCGCGCCGACATG-dt(30)NN-3'. Full-length
                enriched library was constructed using the Clontech
                Creator SMART kit and size-selected to contain the 0.5 kb
                size fraction. Library created in the laboratory of M.
                Brownstein (NIH, NIH). Note: this is a NIH_MGC Library."
BASE COUNT  151 a 205 c 201 g 158 t
ORIGIN

Alignment Scores:
Pred. No.:      6,46e-61      Length:      715
Score:          823.00      Matches:    153
Percent Similarity: 92.90%      Conservative: 17
Best Local Similarity: 83.61%      Mismatches: 13
Query Match:    66.42%      Indels:    0
DB:             14      Gaps:      0

US-09-938-391-2 (1-230) x CB596713 (1-715)

QY      48  ThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnProGly 67
Db      5  ACTCATCAGGACTTTTCAGCGAGTGTCTCCACCTGGTGGCCTCAACACCCCTGTCGGA 64
QY      68  GlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnAlaAlaArgAlaGly 87
Db      65  GGCATCGGTGGTATCCGTGGAGCAGAGTTTCCAGTGTCTTCAGCAGCCCGAGCCGCGGG 124
QY      88  LeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIleVal 107
Db      125  CTGTCCGGCACCTTCGCGGCTTCTCTGTCTCTAGCTCAGGATCTCTATAGCATCGTG 184
QY      108  ArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhePro 127
Db      185  CGCCGCGTGAAGCGGGGGTGTGTGCCCTCGCAACCTGAAGGAGGAGGTGTATCTCCC 244
QY      128  SerTrpGluAlaLeuPheSerGlySerGlyGluGlyGlnLeuLeuPheProGlyAlaArgIlePhe 147

```

```

Db      245  AGCTGGGACTCCTGTTTCTGGCTCCAGGGTCAACTGCAACCCCGGGCCCGCATCTTT 304
QY      148  SerPheAspGlyArgAspValLeuGlnHisProAlaIleTrpProArgLysSerValTrpHis 167
Db      305  TCTTTTGACGGCAGAGATGCTCTGAGACACCCAGCAGCTGGCCGCGAGAGCGTATGGCAC 364
QY      168  GlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThrGlu 187
Db      365  GGCTCGGACCCAGTGGCGGAGGCTGATGGAGATTACTGTGAGACATGGCGCACTGAA 424
QY      188  AlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGlnGlu 207
Db      425  ACTACTGGGGCTACAGGTCAGGCTCCTCCTGCTGTCTCAGCGAGGCTCCTCGGAACAGAAA 484
QY      208  AlaAlaSerCysArgHisAlaPheValValLeuValCysIleGluAsnSerValMetThrSer 227
Db      485  GCTGGAGCTGCCAACACAGCTACATCGTCTGTGCAITTGAGATAGCTTCATGACCTCT 544
QY      228  PheSerLys 230
Db      545  TTCTCCAA 553

RESULT 15
BF074459
LOCUS      221883 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BF074459
ACCESSION  BF074459
VERSION     BF074459.1  GI:10867970
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovinae; Bovinae; Bos.
            1 (bases 1 to 551)
REFERENCE   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrrenkrug,S.C., Bennett
            ,G.L., Heaton,M.P., Laegreid,W., Rohrer,G.A., Chitko-McKown,C.G.,
            Pertea,G., Holt,L., Karamycheva,S., Liang,F., Quackenbush,J. and
            Keele,J.W.
            Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
            Genome Res. 11 (4), 626-630 (2001)
JOURNAL     21180013
MEDLINE     11282978
PUBMED
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smithemail.marc.usda.gov
            Single pass sequencing. Bases called and alt trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCAGTCAGCAGC
Plate: 81 row: F column: 9
Seq primer: ATTTAGTGACACTATAG.
            Location/Qualifiers
            1..551
                /organism="Bos taurus"
                /mol_type="mRNA"
                /db_xref="taxon:9913"
                /tissue_type="pooled"
                /lab_host="DH10B"
                /clone_lib="MARC 2BOV"
                /note="Vector: pCMV SPOR6; Site 1: NotI; Site 2: SalI;
                Library made from pooled tissue from testis, thymus,
                semitendinosus muscle, longissimus muscle, pancreas,
                adrenal, and endometrium."
BASE COUNT  80 a 200 c 179 g 92 t

```


Percent Similarity: 86.19%
Best Local Similarity: 79.52%
Query Match: 68.40%
DB: 12
Conservative: 14
Mismatches: 21
Indels: 8
Gaps: 2

US-09-938-391-2 (1-230) x BG967333 (1-747)

```
QY 1 ProTArpAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
DB 132 CCTGGCGGAGCATGACATCTCTGGCCACCCACCGCGCTGCCAGCCGCGCCTTAC 191
QY 21 ProGlyAlaPro---HisHisGlySerTyrValHisPheGlnProAlaArgProThrGly 39
DB 192 CCGGAGTTCCACATCACCACATCTCTATGTGCACCTCGCGGCGCCGCCGCCACC--- 248
QY 40 GlyProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuVal 59
DB 249 -----CTCTCATCTGTCATCTCATCAGGATCTTCAGCCAGTGTCCACCTGGTG 299
QY 60 AlaLeuAsnSerProGlnProGlyGlyMetArgGlyLeuAlaAspPheGlnCys 79
DB 300 GCACTGAACACCCCTGTCTGGAGGATGCGTGTATCCGTGGAGCAGATTTCCAGTGC 359
QY 80 PheGlnAlaAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerArg 99
DB 360 TTCAGCAGACCGGAGCGGTGGGCTGTCTGGGACCTTCCGGGCTTCTCTCTCTAGG 419
QY 100 LeuGlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValAla 119
DB 420 CTGAGGATCTCTATAGATCTGTCGCGCGTGTGACCGGGGTCTGTGCCATCTGCAAC 479
QY 120 LeuArgAspGluValLeuPheProSerProTyrGluAlaLeuPheSerGlyGln 139
DB 480 CTGAGGACAGGTGTATCTCCAGCTGGGATCTCTGTTTCTGGCTCCAGGGTCAA 539
QY 140 LeuLysProGlyAlaAlaGlyPheSerPheAspGlyArgAspValLeuGlnHisProAla 159
DB 540 CTGCAACCCGGGCGCGCATCTTTCTTTTGAAGCGGAGAGATGCTCTGAGACACCCAGCC 599
QY 160 TrpProArgLysSerValTrpHisGlySerAspProSer-GlyArgArg-LeuThrAspS 179
DB 600 TGCCCGCAGAAAGCGGTATGGCAGCGCTCGGACCCAGTTGGCGGAGGCTGATGGAGA 659
QY 179 erTyrCysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeu- 198
DB 660 GTTACTGTGACATGCGGAACTGAACTACTGGGGCTACAGTCTCTCTCTCTCTCTGG 719
QY 199 LeuAlaGlyArgLeuLeuGln 206
DB 720 CTGTGAGGAGGCTCTCTGGAACAG 743
```

RESULT 13
BE908201
LOCUS 707 bp mRNA linear EST 20-OCT-2000
DEFINITION 601500458F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3902175 5', mRNA sequence.

ACCESSION BE908201
VERSION BE908201.1 GI:10402537
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 707)
NIH-MGC <http://mgi.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLN9704 row: d column: 16
High quality sequence stop: 688.
Location/Qualifiers

FEATURES
source

1..707
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3902175"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 70"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.1 kb. Library constructed by Life Technologies."
BASE COUNT 113 a 250 c 225 g 119 t

ORIGIN

Alignment Scores:
Pred. No.: 5,938-62 Length: 707
Score: 835.00 Matches: 154
Percent Similarity: 92.8% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 67.39% Indels: 0
DB: 10 Gaps: 0

US-09-938-391-2 (1-230) x BE908201 (1-707)

```
QY 47 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 66
DB 11 CACAGCCACCGGAGCTTCCAGCGGTGTCTCCACCTGGTGGCTCAACAGCCCCCTGTCA 70
QY 67 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnAlaArgAlaAla 86
DB 71 GCGCGCATCGCGGGCATCGCGGGCGGCGGCTTCCAGTGTCTCCAGCAGCGGGCGGTG 130
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
DB 131 GGGCTGGCGGGGACCTTCCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 190
QY 107 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 126
DB 191 GTGGCGGTGGCGGACCGCGCGCGCTGCTCCATCGTCAACCTCAAGGACGAGCTGTGTT 250
QY 127 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 146
DB 251 CCGAGCTGGGAGGCTCTGTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGGCGCAGCATC 310
QY 147 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 166
DB 311 TTCTCTTTTGGCGGAAGGACGCTCTGAGGACCCCACTTGGCCCCCAGAGGAGCTGTGG 370
QY 167 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 186
DB 371 CATGGCTCGGACCCCAACCGGCGGAGGCTGACGAGAGCTACTGTGAGACGTGGCGGACG 430
QY 187 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGln 206
DB 431 GAGGCTCTCTCGGACCGGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 490
QY 207 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 226
DB 491 AGTGCCGCGAGGTGCCATCAGCGCTTACATCGTCTCTGATTGAGAACAGCTTCATGACT 550
QY 227 Ser 227
DB 551 GCC 553
```

RESULT 14
CB596713

```

|||||
535 GGAACAGAGAGCTCGAGCTGCCACACAGCTACATGCTGTGCTTGGATTGAGATAGCTTC 594
|||||
225 MetThrSerPheSer 229
|||||
595 ATGACCTCTTTCTCC 609

RESULT 11
BE906253
LOCUS
DEFINITION
757 bp mRNA linear EST 20-OCT-2000
601502237F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3904208 5',
mRNA sequence.
ACCESSION
BE906253
VERSION
BE906253.1 GI:10399595
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 757)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
JOURNAL
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9709 row: i column: 09
High quality sequence stop: 757.
Location/Qualifiers
1..757
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3904208"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 70"
/notes="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
BASE COUNT 145 a 258 c 235 g 119 t
ORIGIN

Alignment Scores:
Pred. No.: 8,36e-64 Length: 757
Score: 857.00 Matches: 160
Percent Similarity: 89.74% Conservative: 15
Best Local Similarity: 82.05% Mismatches: 16
Query Match: 69.17% Indels: 4
DB: 10 Gaps: 1

US-09-938-391-2 (1-230) x BE906253 (1-757)

QY 33 GlnProAlaArgProThrGlyGlyProValHisThrHisThrHisGlnAepPhe 52
|||||
6 CGGCCGGCAGCAGCCACAGCCACCC-----GCCACAGCAGCCGAGCTTC 53
|||||
53 GlnLeuValHisLeuValAlaLeuAenSerProGlnProGlyGlyMetArgGlyIle 72
|||||
54 CAGCGGTGTCTCCACCTGTGTGGCTCAACAGCCCTCTCAGCGGCGCATCGGGGCATC 113
|||||
73 ArgGlyAlaAspPheGlnCysPheGlnAlaArgAlaAlaGlyLeuAlaGlyThrPhe 92
|||||
114 CGCGGGGGCGACTTCAGAGCTTCCAGAGCGCGCGCGGTGGCGTGGCGACCTTC 173
|||||
93 ArgAlaPheLeuSerSerArgLeuGlnAenLeuTySerIleValArgAlaAepArg 112
|||||
174 CGCGCTTCTCTCTCGCGCTGCAGGACCTGTACAGCATGTCGCGCTGCCACCGC 233
|||||
113 ThrGlyValProValValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeu 132
|||||
234 GCAGCGGTGCCCATCGTCAACCTCAAGGACGAGCTGTGTTTCCAGCTGGGAGGCTCTG 293
|||||
133 PheSerGlySerGluGlyGlnLeuLysProGlyAlaArgIlePheSerPheAspGlyArg 152
|||||
294 TTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGGACGCGATCTTCTCTTTCACGGCAG 353
|||||
153 AspValLeuGlnHisProAlaTrpProArgLysSerValTrpHisGlySerAspProSer 172
|||||
354 GACGTCTCTGAGCACCCACCTGGCCAGAGAGCGTGTGCGCATGGCTCGCACCCCAAC 413
|||||
173 GlyArgArgLeuThrAspSerTyrcysGluThrTrpArgThrGluAlaProAlaThr 192
|||||
414 GGGCCAGGCTGACGAGAGCTACTGTGAGCGTGGCGAGCGGCTCCCTCGGCCACG 473
|||||
193 GlyGlnAlaSerSerLeuAlaGlyArgLeuGluGlnGluAlaAlaSerCysArg 212
|||||
474 GGCAGGGCTCTCTCGCTGCTGGGGGCGAGGCTCTGGGGCAGAGTCCCGGAGCTGCCAT 533
|||||
213 HisAlaPheValValLeuCysIleGluAenSerValMetThrSer 227
|||||
534 CACGCTATACGTCTCTGCATTGAGAACAGCTTCATGACTGCC 578
|||||

RESULT 12
BG967333
LOCUS
DEFINITION
747 bp mRNA linear EST 12-JUN-2001
602833649F1 NCI_CGAP_C024 Mus musculus cDNA clone IMAGE:4988053 5',
mRNA sequence.
ACCESSION
BG967333
VERSION
BG967333.1 GI:14354970
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 747)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
JOURNAL
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11000 row: a column: 14
High quality sequence stop: 742.
Location/Qualifiers
1..747
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:4988053"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP C024"
/notes="Organ: Colon; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."
BASE COUNT 135 a 251 c 213 g 148 t
ORIGIN

Alignment Scores:
Pred. No.: 5,38e-63 Length: 747
Score: 847.50 Matches: 167

```

DB: 9 Gaps: 1

US-09-938-391-2 (1-230) x AU125614 (1-715)

QY 1 ProTTPArgAlaAspAspLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
 |||||
 Db 99 CCTGGCGGCAGATGACATCTCTGGCGCAGCCCTCGCTCGCCGAGCCCGCCCTAC 158
 |||||

QY 21 ProGlyValaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
 |||||
 Db 159 CCGGAGCCCCGACACAGCTCTAGTGACCTGGGCGCGCAGACCCACACGCCCA 218
 |||||

QY 41 ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuAlaAla 60
 |||||
 Db 219 CCC-----GCCACAGCCACCGGACTTCAGCGCGTGTCCACCTGGTTGG 266
 |||||

QY 61 LeuAsnSerProGlnProGlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPhe 80
 |||||
 Db 267 CTCACAGCCCTGTGTCAGCGGCATCGCGGCATCGCGCGCGCTTCACCTGCTTC 326
 |||||

QY 81 GlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100
 |||||
 Db 327 CACAGCGCGCGCGTGGGCGCTGGCGGCACCTTCGCGCCCTTCCTGCTCGCGCTG 386
 |||||

QY 101 GlnAspLeuTyrSerIleValaArgAlaAspArgThrGlyValProValValAsnLeu 120
 |||||
 Db 387 CAGGACCTGTACAGATCGTGGCGCGTGGCGCGCGCGCGCGCGCTGCTCACTTC 446
 |||||

QY 121 ArgAspGluValLeuPheProSerTTPGluAlaLeuPheSerGlySerGluGlyGlnLeu 140
 |||||
 Db 447 AAGGACGAGCTGTGTTCAGCTGGGAGCTCTGTCTCAGGCTCTGAGGGTCCGCTG 506
 |||||

QY 141 LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTTP 160
 |||||
 Db 507 AAGCCCGCGCGCAGCATCTCTCTTGACGGACAGGACGCTCTGACCGCCCGCCCTGG 566
 |||||

QY 161 ProArgLysSerValTTPHisGlySerAspProSerGlyArgArgLeuThrAspSerTyr 180
 |||||
 Db 567 CCCAGAGAGCGTGTGGCATGCTCGGACCCCAACGGCGCGCAGGCTGACCGAGCTAC 626
 |||||

QY 181 CysGluThrTTPArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla 200
 |||||
 Db 627 TGTGAGACGTGGCGGAGGAGGCTCCCTCGCGCACGGCGCCAGGCTCTCTCGGTTGCTGGG 696
 |||||

QY 201 GlyArgLeuLeu 204
 |||||
 Db 687 GGCAAGCTCTT 698
 |||||

RESULT 10
 AW911243
 LOCUS
 DEFINITION ur83h10.y1 NCI CGAP Mam6 Mus musculus cDNA clone IMAGE:3156931 5' similar to gb:U17546 Mouse mRNA for collagen (MOUSE);, mRNA sequence.

ACCESSION AW911243
 VERSION AW911243.1
 KEYWORDS GI:8076490
 SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 611)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished
 COMMENT Other ESTs: ur83h10.x1
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-x@mail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: image.llnl.gov/image/html/resources.shtml

MGI:1059687
 Seq primer: -40RP from Gibco
 High quality sequence stop: 403.
 Location/Qualifiers
 1.611
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:3156931"
 /sex="female, virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Mam6"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Salt; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 113 a 199 c 166 g 133 t
 ORIGIN

Alignment Scores:
 Pred. No.: 8,79e-65 Length: 611
 Score: 867.00 Matches: 163
 Percent Similarity: 87.32% Conservative: 16
 Best Local Similarity: 79.51% Mismatches: 22
 Query Match: 69.98% Indels: 4
 DB: 10 Gaps: 1

US-09-938-391-2 (1-230) x AW911243 (1-611)

QY 25 HisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGlyProValHisThr 44
 |||||
 Db 7 CACCACAGTTCTATGTGCACCTGCGCGCAGCCCGCCACC-----CTCTCA 54
 |||||

QY 45 HisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerPro 64
 |||||
 Db 55 CTGTCTCATACTCATCAGGACTTTCAGCCAGTGTCTCCACTGTGTGACACACCCCC 114
 |||||

QY 65 GlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArg 84
 |||||
 Db 115 CTGTCTGGAGGAGCATGCTGTATCTCGTGAGCAGATTTCCAGTGTCTCCAGCAGCCGA 174
 |||||

QY 85 AlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyr 104
 |||||
 Db 175 GCGTGGGCGCTGTCGGGCACCTTCCGGGCTTTCTGTCTCTAGGCTGCAGGATCTCTAT 234
 |||||

QY 105 SerIleValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluVal 124
 |||||
 Db 235 AGCATGTCGCGCGTGTGTCACCGGGGCTGTGTGCCCATCGTCAACTGAGGACGAGGTG 294
 |||||

QY 125 LeuPheProSerTTPGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyAla 144
 |||||
 Db 295 CTATCTCCAGCTGGGACCTCCCTGTTTTCTGGTCCAGGCTCACTGCACACCCGGGGCC 354
 |||||

QY 145 ArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTTPProArgLysSer 164
 |||||
 Db 355 CGCATCTTTCTTTTTCACGCGAGAGATGCTCTGAGACACCCAGCTTGGCGCAGAGAGC 414
 |||||

QY 165 ValTTPHisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTTP 184
 |||||
 Db 415 GTATGCGCAGGCTCGGACCCCGAGGCTGATGAGAGGTACTGTGAGACATGG 474
 |||||

QY 185 ArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeu 204
 |||||
 Db 475 CGAACTGAAACTACTCTGGGCTACAGTACGGCTCTCTCTCTGTGTGTCAGCAGGCTCTG 534
 |||||

QY 205 GluGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerVal 224
 |||||

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 832)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-x@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA sequencing by: NIH Intramural Sequencing Center

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: LLCMI306 row: a column: 14

High quality sequence stop: 679.

Location/Qualifiers

FEATURES

source

1..832

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4582933"

/tissue_type="adenocarcinoma cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 15"

/notes="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adapter: GGACGAG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)"

129 a 320 c 265 g 118 t

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 8,68e-68 Length: 832

Score: 904.00 Matches: 178

Percent Similarity: 86.16% Conservative: 15

Best Local Similarity: 79.46% Mismatches: 25

Query Match: 72.96% Indels: 8

DB: 10 Gaps: 1

US-09-938-391-2 (1-230) x BG387051 (1-832)

QY 1 ProTrpArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20

DB 89 CCCTGGCGGCGAGATGATCTCTGGCCAGCCCTCGCTGCCGAGCCCGAGCCCTAC 148

QY 21 ProGlyAlaProHisHisGlySerTyrValHisGlnProAlaArgProThrGlyGly 40

DB 149 CCGGAGCGCGGACACACACTCTAGTGCACCTGCGCGCGCGGACCAAGCCCA- 207

QY 41 ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuValAla 60

DB 208 CCC-----GCCACACCCAGCGAGTTCAGCCGGTGCTCACCTGGTTGG 255

QY 61 LeuAsnSerProGlnProGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPhe 80

DB 256 CTCAACAGCCCTGTGTCAGCGCGCATCGGGGCGATCCGGGGCGCCAGCTTCAGTGCTTC 315

QY 81 GlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeu 100

DB 316 CAGCAGCGCGGCGCGTGGGGCTGGGGCGGACCTTCGCGCCCTTCCTGCTCGCGCTG 375

QY 101 GlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeu 120

DB 376 CAGACTGTACAGATCGTGGCGCGCGGCGGACCGCGAGCCCGTGCCTCACTCACTC 435

QY 121 ArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeu 140

DB 121 ArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeu 140

Db 436 AAGGACGAGCTGCTGTTTCCAGCTGGGAGGCTCTGTTCTCAGGCTCTCAGGGTCCGCTG 495

QY 141 LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTyr 160

Db 496 AAGCCCGGGGACGATCTTCTTTGACGGCAAGGAGCTCTTGAGGACCCACCTGG 555

QY 161 ProArgLysSerValTrpHisGlySerAsp-ProSerGlyArgArgLeuThr-AspSer 180

Db 556 CCCAGACGAGCGTGTGGCATGGCTCGGACCCCAACGGGCGCGAGCTGACCCGAGAGCT 615

QY 180 YrCysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeu 200

Db 616 ACTGTGAGACGTGGCGGACGAGGCTCCCTCGACACCGGCGGCTC-TCGCTGCTGG 674

QY 200 laGlyArgLeuLeuGluGlnAlaAlaSerCysArgHisAlaPheValValLeuCysI 220

Db 675 GGGGAGGCTCTTGGGGCAGAGTGGCGAGAGCTTGCTCAGGCTACATGCTGTGCA 734

QY 220 leGluAsn 222

Db 735 TTGAGAAC 742

RESULT 9

AU125614

LOCUS

DEFINITION

715 bp mRNA linear EST 01-AUG-2002

AU125614 NT2RM4 Homo sapiens cDNA clone NT2RM4001897 5', mRNA

sequence.

ACCESSION

AU125614

VERSION

AU125614.1 GI:10950330

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 715)

Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,

Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and

Isogai,T.

HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,

Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki

,Y., Sugano,S., Isogai,T.)

Unpublished

Contact: Takao Isogai

Genomics Laboratory

Helix Research Institute

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: genomics@hri.co.jp

HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix

Research Institute; cDNA library construction: Department of

Virology, Institute of Medical Science, University of Tokyo, and

Helix Research Institute.

Location/Qualifiers

1..715

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="NT2RM4001897"

/cell_type="teratocarcinoma"

/cell_line="NT2"

/clone_lib="NT2RM4"

/notes="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal

precursor cells"

BASE COUNT 111 a 278 c 220 g 103 t 3 others

ORIGIN

Alignment Scores:

Pred. No.: 1,06e-67 Length: 715

Score: 902.00 Matches: 165

Percent Similarity: 88.24% Conservative: 15

Best Local Similarity: 80.88% Mismatches: 20

Query Match: 72.80% Indels: 4

```

Db 292 GCACCTGAACACCCCTCTGCGAGCAGTGCCTGGTATCCGTGGAGCAGATTCCAGTGC 351
Qy 80 PheGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerArg 99
Db 352 TTCACGACACCCGAGCCGCGGGCTGTGGGACCTTCGGGCTTCCTGCTCTAG 411
Qy 100 LeuGlnAspLeuTyrSerIleValArgAlaAspArgThrGlyValProValValAsn 119
Db 412 CTGCAGATCTCTATAGCATCGTGCCTGCTGACCGGGGCTGTGCCCATCGTCAAC 471
Qy 120 LeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGln 139
Db 472 CTGAGGACAGGAGTGTATCTCCAGCTGGACTCCCTGTTTCTGGCTCCAGGCTCAA 531
Qy 140 LeuLysProGlyAlaArgIlePheSerPheAspArgAspValLeuGlnHisProAla 159
Db 532 CTGCACCCCGGGCCCGCATCTTTCTTTGACGGCAGAGATGCTCTGAGACACCCAGCC 591
Qy 160 TrpProArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSer 179
Db 592 TGGCCGACAGAGCGTATGGCAGCGCTCGGACCCAGTGGGCGGAGGCTGATGGAGAT 651
Qy 180 TyrCysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeu 199
Db 652 TACTGTGAGACATGGCGAACTGAACTACTGGGGCTACAGTCCAGGCTCTCTCCCTGCTG 711
Qy 200 AlaGlyArg---LeuLeuGluGlnGluAlaAlaSerCysArgHis-AlaPheValValle 218
Db 712 TCAGGACGAGGCTCCCTGGGAACAGAGCTGGCAGCTGCCACACAGCTTACATCGTCT 771
Qy 218 uCysIle-GluAsnSerValMetThrSer-PheSerLys 230
Db 772 GTGCATTGGAGATAGCTTCATGACCTCTATTCTCCAAA 810

RESULT 7
BUS56872
LOCUS
DEFINITION
AGENCY 10186760 NIH_MGC_109 Homo sapiens cDNA clone
IMAGE:6584644 5', mRNA sequence.
BUS56872
VERSION
KEYWORDS
SOURCE
EST.
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 947)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1CM2795 row: f column: 04
High quality sequence stop: 616.
Location/Qualifiers
1..947
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6584644"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_109"
/notes="Organ: ovary; Vector: pOT7; Site: 1; EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned

```

FEATURES source

```

RESULT 8
BUS387051
LOCUS
DEFINITION
602454749F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4582933 5',
mRNA sequence.
BUS387051
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Homo sapiens

```

```

BASE COUNT 163 a 353 c 311 g 120 t
ORIGIN
Alignment Scores:
Pred. No.: 947
Score: 906.00 Matches: 168
Percent Similarity: 89.66% Conservative: 14
Best Local Similarity: 82.76% Mismatches: 16
Query Match: 73.12% Indels: 5
DB: 13 Gaps: 1

```

US-09-938-391-2 (1-230) x BUS56872 (1-947)

```

Qy 1 ProTrpArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
Db 136 CCTGGCGGGCAGATGACATCTGCGCAGCCCTCGCTGCGCGAGCCCGAGCCTAC 195
Qy 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
Db 196 CCGGAGCCCGCAGCAGCTCTACGTGCACCTGCGCGCGGCGCAGCCACACAGCCCA 255
Qy 41 ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuValAla 60
Db 256 CCC-----GCCACAGACACCGCGACTTCCAGCGGTGCTCCACCTGGTGG 303
Qy 61 LeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPhe 80
Db 304 CTCACAGCCCTGTGTCAGCGCGCATGCGGGGATCCGCGGGGCGGACTTCCAGTGTCT 363
Qy 81 GlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100
Db 364 CAGCAGCGCGGGCGCTGGGGCTGGCGGCGACCTTCGCGGCTTCTGCTCGCGCTG 423
Qy 101 GlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeu 120
Db 424 CAGGACCTGTACAGCATGTCGCGCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 483
Qy 121 ArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlnLeu 140
Db 484 AAGGACGAGCTGCTGTTTCCAGCTGGAGGCTCTGTTCTCAGGCTCTGAGGGTCCGCTG 543
Qy 141 LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTrp 160
Db 544 AAGCCCGGGCGCGCATCTTCTCTTTCAGCGCAGGAGCTCTCTGAGGCGGCGGCGGCGG 603
Qy 161 ProArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTyr 180
Db 604 CCCAGAGAGCGGTGTGGCATGCTCGGACCCACAGCGGCGGAGGCTGACGAGAGCTAC 663
Qy 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGln-AlaSerSerLeuLeuAl 200
Db 664 TGTGACGCTGGCGGAGGAGGCTCTCTCGGCCACCGGGGCGGCGGCGGCGGCGGCGG 723
Qy 200 aGlyArg 202
Db 724 GGGCGAG 730

```

```

RESULT 8
BUS387051
LOCUS
DEFINITION
602454749F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4582933 5',
mRNA sequence.
BUS387051
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Homo sapiens

```

into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."


```

Db      749 ATTGAGACCGCTTCATGACTGCC 772
|||||||
RESULT 4
BI412588      874 bp      mRNA      linear      EST 14-AUG-2001
LOCUS      602990468F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5146409 5',
DEFINITION      mRNA sequence.
ACCESSION      BI412588
VERSION      BI412588.1 GI:15173511
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 874)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-x@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLML1361 row: g column: 18
High quality sequence start: 23
High quality sequence stop: 808.
FEATURES
source
1..874
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:5146409"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NCI_CGAP_Lu33"
notes="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'.
TGTACCACTCAAGTGGAGCGCGCTCTGTTTTTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      189 a      275 c      228 g      181 t      1 others
ORIGIN
Alignment Scores:
Pred. No.:      1..14e-71      Length:      874
Score:      949.50      Matches:      183
Percent Similarity:      87.12%      Conservative:      20
Best Local Similarity:      78.54%      Mismatches:      23
Query Match:      76.63%      Indels:      7
DB:      12      Gaps:      2
US-09-938-391-2 (1-230) x BI412588 (1-874)
Qy      1 ProTprArg-AlaAspAspIleuAlaGlyProProArgLeuLeuAspProGlnProTy 20
Db      12 CCCTGGCGATTTCAGATGACATCTCTGGCCAAACCCACCGCGCTGCGACGCCGACCTTA 71
Qy      20 rProGlyAla-Pro---HisHisGlySerTyValHisPheGlnProAlaArgProThrG 39
Db      72 CCCCAGGATTCCATCACCACAGTTCTTATGTGCACCTGGCCGCGACCGCCGCCACC- 130
39 lyGlyProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuV 59
131 -----CTCTCACTTGTCTCATCTCATCAGGACTTTTTCAGCAGTGTCTCCACCTGG 179
59 alAlaLeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAappPheGlnC 79
180 TGGCACTGAACACCCCTGTCTGGAGGATGCGTGGTATCCGTGGAGCAGATTTCAGT 239
79 yspPheGlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSera 99
240 GCTTCAGCAAGCCGAGCCGTGGGCTGTGGGACCTTCCGGGCTTTCCTGTCTCTA 299
99 rgLeuGlnAspLeuTySerIleValArgArgAlaAspArgThrGlyValProValValA 119
300 GGCTGCAGGATCTTATAGCATCGTGGCGCTGTGGGACCTTCCGGGCTTTCCTGTCTCTA 359
119 snLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyG 139
360 ACCTGAAGGACGAGGTGCTATCTCCAGCTGGAGTCCCTGTTTCTGGCTCCCGGGTC 419
139 lnLeuLysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProA 159
420 AACTGCAACCCGGGCGCGCATCTTTCTTTTGACGGCAGAGATGCTCTGACACACCCAG 479
159 laTrpProArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspS 179
480 CTGGCGCGCAGAAAGCGTATGCGACGGCTCGGACCCAGTGGCGGAGGCTGATGAGAGA 539
179 erTyrcysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuL 199
540 GTTACTGTGAGACATGGCGAACTGAACACTACTGGGGCTACAGGTACAGGCTCCTCCCTGC 599
199 euAlaGlyArgLeuLeuGluGlnGluAlaAlaSerCysArgHisAlaPheValValLeuC 219
600 TGTACGAGGCTCTCTGGACAGAAAGTCCGAGTCCAGGCTCCACACAGCTACATGCTCTGT 659
219 ysIleGluAsnSerValMetThrSerPheSerLys 230
660 GCATTGAGAATAGTTCATGACCTCTTTCTCCAAA 694
RESULT 5
BQ673186
LOCUS
DEFINITION      AGENCOURT_8414390 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6272287
5', mRNA sequence.
ACCESSION      BQ673186
VERSION      BQ673186.1 GI:21784020
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 979)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-x@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2450 row: g column: 08
High quality sequence stop: 638.
FEATURES
source
1..979
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6272287"

```



```
QY 21 ProGlyAlaPro--HisHisGlySerTyrValHisPheGlnProAlaArgProThrGly 39
Db 102 CCGGAGTTCCACATCACCACAGTCTCTATGTGCACCTGCGCGCACCGCCGCCACCC---158
QY 40 GlyProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValHisLeuVal 59
Db 159 -----CTCCACCTGCTCATCTACTCATCAGGACTTTTCAGCCAGTGTCTCCACCTGGTG 209
QY 60 AlaLeuAsnSerProGlnProGlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCys 79
Db 210 GCACCTGAACACCCCTCTCTGAGGAGTCTGCTGGATCGTGGAGACAGATTTCCAGTGC 269
QY 80 PheGlnGlnAlaArgAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArg 99
Db 270 TTCCAGCAAGCCCGAGCGGTGGGGCTGTCGGGCACCTTCGGGGCTTCTGTCTCTAGG 329
QY 100 LeuGlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsn 119
Db 330 CTGCAGGATCTCTATAGCATCTGCGCCGCTGCTGACCGGGGTCTGTGCCCATCTGTCAC 389
QY 120 LeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGln 139
Db 390 CTGAAGGACAGAGTGTATCTCCAGCTGGGACTCCCTCTGTTTCTGGCTCCCGAGGTCAA 449
QY 140 LeuLysProGlyAlaAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAla 159
Db 450 CTGCAACCCGGGGCCGCATCTTTCTTTGACGGCAGAGATGCTCTGAGACACCCAGCC 509
QY 160 TrpProArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSer 179
Db 510 TGGCCCGCAGAAAGCGTATGGACGCGCTCGGACCCAGTGGCGGAGGTGTATGGAGAGT 569
QY 180 TyrCysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeu 199
Db 570 TACTGTGAGACATGGCGAACTGAACTACTGGGGCTACAGGTGAGGCTCTCTCCCTGCTG 629
QY 200 AlaGlyArgLeuLeuGluGlnAlaAlaSerCysArgHisAlaPheValValLeuCys 219
Db 630 TCAGGCGAGCTCTCGAAGAGAGTGGAGGTGCGACACACAGCTACATCTGCTCTGTCG 689
QY 220 IleGluAsnSerValMetThrSerPheSerLys 230
Db 690 ATTGAGATAGCTTTCATGACCTCTTCTCCAAA 722

RESULT 3
BQ672290
LOCUS
DEFINITION BQ672290 929 bp mRNA linear EST 15-JUL-2002
5', mRNA sequence.
AGENCOURT 8354946 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6275261
ACCESSION BQ672290
VERSION BQ672290.1 GI:21783124
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 929)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2458 row: c column: 06
High quality sequence stop: 588.
Location/Qualifiers
1..929
```

FEATURES
source

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6275261"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_102"
/Note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-CDNA
synthesis kit (stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 152 a 373 c 272 g 132 t
ORIGIN
Alignment Scores:
Pred. No.: 1,71e-72 Length: 929
Score: 959.50 Matches: 179
Percent Similarity: 86.84% Conservatives: 19
Best Local Similarity: 78.51% Mismatches: 25
Query Match: 77.44% Indels: 5
DB: 13 Gaps: 2
US-09-938-391-2 (1-230) x BQ672290 (1-929)
```

```
QY 1 ProTrpArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
Db 101 CCTCGCGCGGAGATGATCATCTTGGCCAGACCCCTCGCTCGCCGAGCCCGCCCTAC 160
QY 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
Db 161 CCCGAGAGCCCGCACACACAGCTCTTACGTGACCTTCCGCGCGCGCGAGCCCAAGCCCA 220
QY 41 ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuValAla 60
Db 221 CCC-----GCCACAGCCACCGGACTTCCAGCCGCTGCTCCACCTGGTGGG 268
QY 61 LeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPhe 80
Db 269 CTCACAGACCCCTGTCAGCGGCGCATGCGGGGCATCCGCGGCGCGACTTCACAGTGTCTC 328
QY 81 GlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100
Db 329 CAGCAGGCGCGGCGCTGCGGGCTGCGGGCACCTTCCGCGCTTCTGTCTCGCGCCTG 388
QY 101 GlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeu 120
Db 389 CAGGACCTGTACAGCATCTGTGCGCGCTGCCGACCGCGGAGCGCTGCCCATCTCAACCTC 448
QY 121 ArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeu 140
Db 449 AAGGACGAGCTGCTGTTCCAGCTGGGAGGCTCTGTTCTCAGGCTCTGAGGCTCGCGCTG 508
QY 141 LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTrp 160
Db 509 AAGCCCGGCGGACGATCTTCTCTTAAACGCAAGGACGCTCTGAGCAGCCACCTCG 568
QY 161 ProArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTyr 180
Db 569 CCCACAGAGAGCTGTGTGATGGCTCGGACCCCAACGCGGCGAGGCTGACCGAGAGTAC 628
QY 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla 200
Db 629 TGTGAGACGTGCGGACGAGGCTCCCTCGGCGACCGGCGAGGCTCTCTCTCTCTCTCTGGG 688
QY 201 GlyArgLeu---LeuGluGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCys 219
Db 689 GGGCAGGCTCCCTGGGCGAGAGTGGCCGCGAGCTCCCATACGCTATCATCTGCTCTGC 748
QY 220 IleGluAsnSerValMetThrSer 227
```


AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: NDAM424 row: g column: 24
 High quality sequence stop: 689.
 Location/Qualifiers
 1..881
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30365831"
 /tissue_type="Pituitary"
 /lab_host="DH108-Ton A (T1 and T5 phage resistances)"
 /clone_lib="NIH_MGC_179"
 /notes="Organ: brain; Vector: pCMV-SPORT6.1; Site 1: EcoRV (destroyed); Site 2: NotI; Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.1 kb. Library was constructed by (Invitrogen). Note: this is a NIH_MGC Library."
 BASE COUNT 138 a 341 c 278 g 124 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.24e-77 Length: 881
 Score: 1019.00 Matches: 187
 Percent Similarity: 89.43% Conservative: 16
 Best Local Similarity: 82.38% Mismatches: 20
 Query Match: 82.24% Indels: 4
 DB: 14 Gaps: 1
 US-09-938-391-2 (1-230) x CD105862 (1-881)
 QY 1 ProTprArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
 Db 166 CCCTGGCGGCGAGATGACATCTCTGGCGAGCCCTCGCTGCGCGAGCCCGAGCCCTAC 225
 QY 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
 Db 226 CCGGAGCCCCCGACACACAGCTCTACGTGCACCTCGCGCGCGCACGACCCACAGCCCA 285
 QY 41 ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuValAla 60
 Db 286 CCC-----GCCACACCCACCGGACTTCCAGCGGTGCTCCACCTGTTGG 333
 QY 61 LeuAsnSerProGlnProGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPhe 80
 Db 334 CTCAACAGCCCCCTGTGAGCGGCGCATCGGGGCATCGCGGGCGCGACTTCCAGTCTTC 393
 QY 81 GlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeu 100
 Db 394 CACAGCGCGCGGCGCGTGGGGCTGGGGGGGACCTTCGCGGCTTCCTGCTCGCCCTG 453
 QY 101 GlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeu 120
 Db 454 CAGGACCTGTACAGATCTGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 513
 QY 121 ArgAspGluValLeuPheProSerTprGluAlaLeuPheSerGlySerGluGlyGlnLeu 140
 Db 514 AAGGACGAGCTGTGTTTCCAGCTGGAGGCTCTGTCTCAGGCTCTGAGGGTCCGCTG 573
 QY 141 LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTpr 160
 Db 574 AAGCCCCGGGCGACGCATCTCTCTTGTGAGCGAAGGAGTCTCTGAGGACCCACCTGG 633

QY 161 ProArgLysSerValTrpHisGlySerAspProSerGlyArgLeuThrAspSerTyr 180
 Db 634 CCCAGAGAGCGTGTGGCATGCTCGGACCCCAACGCGCGCGAGCTGACCGAGAGCTAC 693
 QY 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla 200
 Db 694 TGTGAGACGTGGCGAGCGAGGCTCCCTCGGCCACGGCGCGGCGCTCTCGCTGCTGGG 753
 QY 201 GlyArgLeuLeuGluGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCysIle 220
 Db 754 GCGAGCTCTCTGGGCGAGAGTCCGCGGAGCTGCCATCAGCCCTACATCGTCTGCTGCA 813
 QY 221 GluAsnSerValMetThrSer 227
 Db 814 GAGAACAGCTTCATGACTGCC 834
 RESULT 2
 BF385854
 LOCUS 602046021F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:4195660 5',
 DEFINITION mRNA sequence.
 ACCESSION BF385854
 VERSION BF385854.1 GI:11367159
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE 1 (bases 1 to 843)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM9531 row: i column: 05
 High quality sequence stop: 761.
 Location/Qualifiers
 1..843
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="FVB/N"
 /clone="IMAGE:4195660"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI_CGAP_Li9"
 /note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
 BASE COUNT 167 a 278 c 228 g 170 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.31e-74 Length: 843
 Score: 983.50 Matches: 185
 Percent Similarity: 87.88% Conservative: 18
 Best Local Similarity: 80.09% Mismatches: 23
 Query Match: 79.38% Indels: 5
 DB: 10 Gaps: 2
 US-09-938-391-2 (1-230) x BF385854 (1-843)
 QY 1 ProTprArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
 Db 42 CCCTGGCGGCGAGATGACATCTCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTTAC 101

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 17, 2003, 19:46:50 ; Search time 2038.33 Seconds
(without alignments)
2742.451 Million cell updates/sec

Title: US-09-938-391-2
Perfect score: 1239
Sequence: 1 PWRADILLAPRLLDPQY.....CRHAFVVLCIENSVMTSFSK 230

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DRV=xlpl
-Q=/cgn2_1/USPTO.spool_p/US09938391/runat_04082003_130656_27102/app_query.fasta_1.718
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEARSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938391@cgn_1_1_4399@runat_04082003_130656_27102 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pin.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_pbg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB	ID	Description
1	1019	82.2	881	14	CD105862	AGENCOURT
2	983.5	79.4	843	10	BF385854	BF385854 602046031
3	959.5	77.4	929	13	BQ672290	BQ672290 AGENCOURT
4	949.5	76.6	874	12	BI412588	BI412588 602990468
5	947.5	76.5	979	13	BQ673186	BQ673186 AGENCOURT
6	927	74.8	835	12	BI526580	BI526580 AGENCOURT
7	906	73.1	947	13	BU556872	BU556872 AGENCOURT
8	904	73.0	832	10	BG387051	BG387051 602454749
9	902	72.8	715	9	AU125614	AU125614 AU125614
10	867	70.0	611	10	AW911243	AW911243 ur83h10.Y
11	857	69.2	757	10	BE906253	BE906253 601502237
12	847.5	68.4	747	12	BG67333	BG67333 602833649
13	835	67.4	707	10	BE908201	BE908201 601500458
14	823	66.4	715	14	CB596713	CB596713 AGENCOURT
15	822	66.3	551	10	BF074459	BF074459 221883 MA
16	815.5	65.8	782	9	AI326391	AI326391 mm18f09.X
17	814	65.7	753	12	BI904605	BI904605 603168411
18	809.5	65.3	587	9	AA288198	AA288198 vbl5b10.f
19	805	65.0	897	12	BI080524	BI080524 602877005
20	804	64.9	618	9	AV696242	AV696242 AV696242
21	797.5	64.4	755	12	BG969001	BG969001 602834938
22	794	64.1	733	10	BG072504	BG072504 H311D11-
23	793	64.0	723	12	BI247582	BI247582 602960041
24	784	63.3	944	13	B0859398	B0859398 AGENCOURT
25	781	63.0	915	10	BF166139	BF166139 601776586
26	780.5	63.0	595	9	AA003030	AA003030 mg50606.f
27	772	62.3	720	12	BI147444	BI147444 602914008
28	771	62.2	846	13	BU540812	BU540812 AGENCOURT
29	770.5	62.2	890	13	BQ896475	BQ896475 AGENCOURT
30	770	62.1	703	13	BU615520	BU615520 UI-H-FG0-
31	760	61.3	706	13	BU439577	BU439577 603208252
32	757	61.1	652	13	BU352506	BU352506 603527982
33	753	60.8	683	12	BM683067	BM683067 UI-B-E01-
34	752.5	60.7	536	9	AA277488	AA277488 v882007.f
35	752	60.7	690	12	BI219399	BI219399 602936756
36	750	60.5	795	14	CB601615	CB601615 AGENCOURT
37	744	60.0	657	14	CB444165	CB444165 695295 MA
38	740.5	59.8	585	9	AA212785	AA212785 mw78908.f
39	739	59.6	682	9	AW089583	AW089583 xg20f03.X
40	736.5	59.4	614	13	BU459935	BU459935 603367327
41	735	59.3	884	12	BI161007	BI161007 602865213
42	732	59.1	1093	13	BQ723254	BQ723254 AGENCOURT
43	726	58.6	592	13	BQ567187	BQ567187 g180h10.Y
44	726	58.6	633	11	AK014292	AK014292 Mus muscu
45	726	58.6	653	10	BF384828	BF384828 602046324

ALIGNMENTS

RESULT 1
LOCUS CD105862
DEFINITION AGENCOURT_14021788 NIH_MGC_179 Homo sapiens CDNA clone
IMAGE:30365831 5', mRNA sequence.
ACCESSION CD105862
VERSION CD105862.1 GI:30759036
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 881)

Search completed: August 18, 2003, 00:40:15
Job time : 283.444 secs

/ FILING DATE: <Unknown>
/ APPLICATION NUMBER: US 08/608,845
/ FILING DATE: 16-JUL-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McMorris Jr., Robert G
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (302) 658-9141
/ TELEFAX: (302) 658-5613
/ INFORMATION FOR SEQ ID NO: 37:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 565 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 37:
US-10-036-869-37

Alignment Scores:
Pred. No.: 9,74e-84 Length: 565
Score: 831.00 Matches: 154
Percent Similarity: 92.93% Conservative: 17
Best Local Similarity: 83.70% Mismatches: 13
Query Match: 67.07% Indels: 0
DB: 13 Gaps: 0

US-09-938-391-2 (1-230) x US-10-036-869-37 (1-565)

```
QY 47 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 66
DB 11 CATACTCATCAGGACTTTCAGCGAGTTCCTGCTCTAGGCTGAGGATCTCTATAGCATC 70
QY 67 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAla 86
DB 71 GGAGGATCGTGTATCCGTGAGCAGATTCAGTGTCTTCAGACAGCCGAGCGG 130
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSer 106
DB 131 GGGCTGTCGGGACCTTCGGGCTTTCCTGCTCTAGGCTGAGGATCTCTATAGCATC 190
QY 107 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 126
DB 191 GTCGCGCGTCTGACCGGGGCTGTGCTCCATCTCAACCTGAAGACGAGGTGTATCT 250
QY 127 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArg 146
DB 251 CCCAGCTGGGACTCCCTGTTCTGCTCCAGGCTCAACCTGAGGAGGAGGAGGAGGAG 310
QY 147 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerVal 166
DB 311 TTTTCTTTTTCAGCGCAGAGATGCTCTGAGACACCCAGCCTGGCGCGCAGAGAGCG 370
QY 167 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArg 186
DB 371 CACGGCTCGACCCAGTGGCGGAGGCTGTATGAGAGGTTACTGTGAGACATGTCGCA 430
QY 187 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuAlaGlyValLeuGluGln 206
DB 431 GAACTACTGGGGCTACAGTCAAGGCTCTCTCTGTCAGGAGGCTCTCTGGAGACAG 490
QY 207 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMet 226
DB 491 AAAGCTGCGAGCTGCCACACAGCTACATCTCTCTGTGCTGAGATAGCTTCATGACC 550
QY 227 SerPheSerLys 230
DB 551 TCTTTCTCCAAA 562
```

RESULT 15

US-10-292-418-17
/ Sequence 17, Application US/10292418
/ Publication No. US20030139365A1
/ GENERAL INFORMATION:
/ APPLICANT: Lo, Kin-Ming

/ APPLICANT: Li, Yue
/ APPLICANT: Gillies, Stephen D
/ TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
/ TITLE OF INVENTION: Immunofusins
/ FILE REFERENCE: LEX-006C1
/ CURRENT APPLICATION NUMBER: US/10/292,418
/ CURRENT FILING DATE: 2002-11-12
/ PRIOR APPLICATION NUMBER: 09/383,315
/ PRIOR FILING DATE: 1999-08-25
/ PRIOR APPLICATION NUMBER: US 60/097,883
/ PRIOR FILING DATE: 1998-08-25
/ NUMBER OF SEQ ID NOS: 54
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 17
/ LENGTH: 552
/ TYPE: DNA
/ ORGANISM: Mus musculus
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(552)
/ OTHER INFORMATION: endostatin
US-10-292-418-17

Alignment Scores:
Pred. No.: 2,05e-83 Length: 552
Score: 828.00 Matches: 153
Percent Similarity: 92.93% Conservative: 18
Best Local Similarity: 83.15% Mismatches: 13
Query Match: 66.83% Indels: 0
DB: 12 Gaps: 0

US-09-938-391-2 (1-230) x US-10-292-418-17 (1-552)

```
QY 47 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 66
DB 1 CATACTCATCAGGACTTTCAGCGAGTTCCTGCTCTAGGCTGAGGATCTCTATAGCATC 60
QY 67 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAla 86
DB 61 GGAGGATCGTGTATCCGTGAGCAGATTCCTGCTCTAGGCTGAGGAGGAGGAGGAGG 120
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSer 106
DB 121 GGGCTGTCGGGACCTTCGGGCTTTCCTGCTCTAGGCTGAGGATCTCTATAGCATC 180
QY 107 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 126
DB 181 GTCGCGCGTCTGACCGGGGCTGTGCTCCATCTCAACCTGAAGACGAGGTGTATCT 240
QY 127 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArg 146
DB 241 CCCAGCTGGGACTCCCTGTTCTGCTCCAGGCTCAAGTGCACCCCGGGGCGCGCATC 300
QY 147 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerVal 166
DB 301 TTTTCTTTTTCAGCGCAGAGATGCTCTGAGACACCCAGCCTGGCGCGCAGAGCG 360
QY 167 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArg 186
DB 361 CACGGCTCGACCCAGTGGCGGAGGCTGTATGAGAGGTTACTGTGAGACATGTCGCA 420
QY 187 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuAlaGlyValLeuGluGln 206
DB 421 GAACTACTGGGGCTACAGGCTCAGGCTCCTCTCTGTCAGGAGGCTCTCTGGAGACAG 480
QY 207 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMet 226
DB 481 AAAGCTGCGAGCTGCCACACAGCTACATCTCTCTGTGCTGAGATAGCTTCATGACC 540
QY 227 SerPheSerLys 230
DB 541 TCTTTCTCCAAA 552
```

QY 47 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 66
DB 4 CATACTCATCAGGACTTTTCAGCCAGTGTCTCCACCTGGTGGCATTGAACACCCCTGTCT 63
QY 67 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAla 86
DB 64 GGAGGCATGCGTGGTATCCGTGGAGCAGATTTCAGTGTCTTCAGCAAGCCGAGCGTG 123
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerile 106
DB 124 GGCGCTGTGGGACCTTCGCGGCTTCTCTAGGCTGCAGGATCTCTATAGCATC 183
QY 107 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 126
DB 184 GTGCGCGTGTCTGACCGGGGCTGTGCGCATCGTCAACCTGAAGCAGCGGTGTCTATCT 243
QY 127 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgile 146
DB 244 CCAGCTGGGACTCCCTGTTTCTGGCTCCAGGTCACCTGCAACCCGCGGCGCGCATC 303
QY 147 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTyr 166
DB 304 TTTTCTTTGACGGAGATGCTTGAGACACCCAGCTGGCGCGCAGAGCGTATGG 363
QY 187 GluAlaProAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGln 206
DB 424 GAAACTACTGGGCTACAGTCAAGGCTCTCCCTGTGTGAGGAGCTCTCTGGAACAG 483
QY 207 GluAlaAlaSerCysArgHisAlaPheValValLeuCysLeuGlnAsnSerValMetThr 226
DB 484 AAGTGTGGAGTGGCCACACAGCTACATCGTCTGTGAGGAGCTCTCTGGAACAG 543
QY 227 SerPheSerLys 230
DB 544 TCTTTCTCCAAA 555

RESULT 13

US-09-775-325-3

; Sequence 3, Application US/09775325

; Patent No. US20020076397A1

; GENERAL INFORMATION:

; APPLICANT: Abitbol, Marc

; APPLICANT: Uteza, Yves

; APPLICANT: Menasche, Maurice

; APPLICANT: Bossard, Carine

; APPLICANT: Van Den Bergh, Loic

; APPLICANT: Bonnel, Sebastian

; APPLICANT: Prats, Herve

; APPLICANT: Honiger, Jiri

; APPLICANT: Neuner-Jehle, Martin

; TITLE OF INVENTION: INTRAOCULAR TRANSPLANTATION OF ENCAPSULATED CELLS

; FILE REFERENCE: 8076.202U51

; CURRENT APPLICATION NUMBER: US/09/775,325

; CURRENT FILING DATE: 2001-02-01

; PRIOR APPLICATION NUMBER: US 09/449,293

; PRIOR FILING DATE: 1999-11-24

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 558

; TYPE: DNA

; ORGANISM: Rattus rattus

US-09-775-325-3

Alignment Scores:

Pred. No.: 9.6e-84 Length: 558

Score: 831.00 Matches: 154

Percent Similarity: 92.93% Conservative: 17

Best Local Similarity: 83.70% Mismatches: 13

Query Match: 67.07% Indels: 0
DB: 9 Gaps: 0

US-09-938-391-2 (1-230) x US-09-775-325-3 (1-558)

QY 47 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 66

DB 4 CATACTCATCAGGACTTTTCAGCCAGTGTCTCCACCTGGTGGCATTGAACACCCCTGTCT 63

QY 67 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAla 86

DB 64 GGAGGCATGCGTGGTATCCGTGGAGCAGATTTCAGTGTCTTCAGCAAGCCGAGCGTG 123

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerile 106

DB 124 GGCGCTGTGGGACCTTCGCGGCTTCTCTAGGCTGCAGGATCTCTATAGCATC 183

QY 107 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 126

DB 184 GTGCGCGTGTCTGACCGGGGCTGTGCGCATCGTCAACCTGAAGCAGCGGTGTCTATCT 243

QY 127 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgile 146

DB 244 CCAGCTGGGACTCCCTGTTTCTGGCTCCAGGTCACCTGCAACCCGCGGCGCGCATC 303

QY 147 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTyr 166

DB 304 TTTTCTTTGACGGAGATGCTTGAGACACCCAGCTGGCGCGCAGAGCGTATGG 363

QY 167 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 186

DB 364 CACGGCTCGGACCCCAAGTGGCGGAGCTGTGAGAGTGTACTGTGAGACATGGCGAACT 423

QY 187 GluAlaProAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGln 206

DB 424 GAAACTACTGGGCTACAGTCAAGGCTCTCCCTGTGTGAGGAGCTCTCTGGAACAG 483

QY 207 GluAlaAlaSerCysArgHisAlaPheValValLeuCysLeuGlnAsnSerValMetThr 226

DB 484 AAGTGTGGAGTGGCCACACAGCTACATCGTCTGTGAGGAGCTCTCTGGAACAG 543

QY 227 SerPheSerLys 230

DB 544 TCTTTCTCCAAA 555

RESULT 14

US-10-036-869-37

; Sequence 37, Application US/10036869

; Publication No. US20020151516A1

; GENERAL INFORMATION:

; APPLICANT: Mixson, James A

; TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA

; THERAPY

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Connolly, Bove, Lodge, & Hutz

; STREET: 1220 Market Street, P.O. Box 2207

; CITY: Wilmington

; STATE: Delaware

; COUNTRY: U.S.A.

; ZIP: 19899

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/036,869

; FILING DATE: 29-No. US20020151516A1-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/985,526

Qy	67	GlyGlyMetArgGlyIleArgGlyAlaAlaAspPheGlnCysPheGlnGlnAlaAraGAlaAla	86
Db	61	GGCGGATCGGGGATCGCGGGGCGAGTTCAGTGTCCAGCAGCGCGGGCGGTG	120
Qy	87	GlyLeuAlaClyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle	106
Db	121	GGGCTGGCGGGCACCTTCGCGGCTTCCTGTCTCGCGCTGCAGGACCTGTACAGCATC	180
Qy	107	ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe	126
Db	181	GTGGCGGTGCCGACCGCGGACGCCATCGTCAACCTCAGAGCAGGACTGCTGTTT	240
Qy	127	ProSerTrrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyValaAraGile	146
Db	241	CCACGTGGGAGGCTCTGTTCTCAGGCTCTGAGGTCGCTGAGGCCGCGGGCAGCATC	300
Qy	147	PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrrpProArgLysSerValTrrp	166
Db	301	TTCTCTTTGACGGCAAGACGTCCTGAGGCACCCACCTGGCCCCGAGAGCGTGTGG	360
Qy	167	HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrrpArgThr	186
Db	361	CATGGCTCGACCCCAACGGCGCAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGACG	420
Qy	187	GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln	206
Db	421	GAGGTCCTTCGGCCACGGCCAGGCTCTCGTGTGGGGGCGAGGCTCTCTGGGCGAG	480
Qy	207	GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr	226
Db	481	AGTGCGCGGAGCTGCCATCGCCTACATCTCGTGTCTGCAATTGAGACAGCTTTCAGCT	540
Qy	227	Ser 227	
Db	541	GCC 543	

RESULT 11
 US-10-131-241-48
 ; Sequence 48, Application US/10131241
 ; Publication No. US20030012792A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Holaday, John W.
 ; APPLICANT: Fortier, Anne H.
 ; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation and Angiogenesis Using Cancer Markers
 ; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
 ; FILE REFERENCE: 05213-0344 43170-271565
 ; CURRENT APPLICATION NUMBER: US/10/131,241
 ; CURRENT FILING DATE: 2002-07-22
 ; PRIOR APPLICATION NUMBER: US 09/413,049
 ; PRIOR FILING DATE: 1999-10-06
 ; PRIOR APPLICATION NUMBER: US 09/316,802
 ; PRIOR FILING DATE: 1999-05-21
 ; PRIOR APPLICATION NUMBER: US 60/086,586
 ; PRIOR FILING DATE: 1998-05-22
 ; NUMBER OF SEQ ID NOS: 65
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 48
 ; LENGTH: 540
 ; TYPE: DNA
 ; ORGANISM: Rhesus monkey
 ; US-10-131-241-48

Alignment Scores:			
Pred. No.:	9,22e-84	Length:	540
Score:	831.00	Matches:	152
Percent Similarity:	92.22%	Conservative:	14
Best Local Similarity:	84.44%	Mismatches:	14
Query Match:	67.07%	Indels:	0
DB:	14	Gaps:	0
US-09-938-391-2 (1-230) x US-10-131-241-48 (1-540)			
Qy	47	HisThrHisGlnAspPheGlnValLeuValAlaLeuAsnSerProGlnPro	66

Db	1	CACAGCACCCGCGACTTCCAGCCCGGTGCTCCACCTGGTTCGCTCAATAGCCCGCTGCCA	60
Qy	67	GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnAlaArgAlaAa	86
Db	61	GGCGGCATCGGGGGATCCGCGGGCCGACTTCAGTGCTTCAGCAGCAGCGGCCGTG	120
Qy	87	GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle	106
Db	121	GGGCTGGTGGGCACCTTCGTTGCCTTCCTCAGCGCTGCAGGACCTGTACAGCATC	180
Qy	107	ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe	126
Db	181	GTGGCGCGTGCAGCCGCGCAGCGTGCCTCAGCTCGTCAACCTCAAGGATGAGCTGCTTT	240
Qy	127	ProSerTrpGluAlaLeuPheSerGlySerGlyGlyGlnLeuLeuYspProGlyAlaArgIle	146
Db	241	CCAGCTGGAGGCTTTGTTGCGAGGCTCTGAGGGTCGCTGAAGCCCGGCGCAGCATC	300
Qy	147	PheSerPheAspGlyArgAspValLeuGlnHisProAlaTatProArgLysSerValTrp	166
Db	301	TTCTCTTTTACGGCAAGGACGTCTCGAGGACCCCACTGGCCCCAGAAAGACGCTGTGG	360
Qy	167	HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr	186
Db	361	CATGGCTCGGACCCAGCGCGCGAGGCTGACTGAGAGCTACTGCGAGACGTGGCGGACA	420
Qy	187	GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGluGln	206
Db	421	GAGTCTCCCTCGGTACAGGTCAGGCTCCTCCCTGCTGGGGGGCAGGCTCTCAGGCGAG	480
Qy	207	GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr	226
Db	481	AATCGCGAAGCTGTCACCGCCTATATCGTCTCTGATCGATCGAAGACAGTTCATGACT	540

RESULT 12
US-09-775-174-3
; Sequence 3, Application US/09775174
; Patent No. US20020076396A1
; GENERAL INFORMATION:
; APPLICANT: Abitbol, Marc
; APPLICANT: Uteza, Yves
; APPLICANT: Menasche, Maurice
; APPLICANT: Bossard, Carine
; APPLICANT: Van Den Bergh, Loic
; APPLICANT: Bonnel, Sebastian
; APPLICANT: Prats, Hervé
; APPLICANT: Honiger, Jiri
; APPLICANT: Neuner-Jehle, Martin
; TITLE OF INVENTION: INTRAOCULAR TRANSPLANTATION OF ENCAPSULATED CELLS
; FILE REFERENCE: 8076.202USD2
; CURRENT APPLICATION NUMBER: US/09/775,174
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: US 09/449,293
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Rattus rattus
; US-09-775-174-3

RESULT 12
US-09-775-174-3
; Sequence 3, Application US/09775174
; Patent No. US20020076396A1
; GENERAL INFORMATION:
; APPLICANT: Abitbol, Marc
; APPLICANT: Uteza, Yves
; APPLICANT: Menasche, Maurice
; APPLICANT: Bosseard, Carine
; APPLICANT: Van Den Bergh, Loic
; APPLICANT: Bonnel, Sebastian
; APPLICANT: Prats, Hervé
; APPLICANT: Honiger, Jiri
; APPLICANT: Neuner-Jehle, Martin
; TITLE OF INVENTION: INTRAOCULAR TRANSPLANTATION OF ENCAPSULATED CELLS
; FILE REFERENCE: 8076.202USD2
; CURRENT APPLICATION NUMBER: US/09/775,174
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: US 09/449,293
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Rattus rattus
US-09-775-174-3

Alignment Scores:		
Pred. No.:	9.6e-84	Length:
Score:	831.00	Matches:
Percent Similarity:	92.93%	Mismatches:
Best Local Similarity:	83.70%	Indels:
Query Match:	67.07%	Gaps:
DB:	9	0
US-09-938-391-2 (1-230) x US-09-775-174-3 (1-558)		

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 121 GGCGTGGGGGACCTTCGCGGCTTCTCTCGCGCTGAGGACCTGTACAGCATC 180
QY 107 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 126
Db 181 GTGCGCGTGGCCGACCGCGCGCTGCTCAAGCTCAAGGACGAGCTGTGTTT 240
QY 127 ProSerTrpGluAlaLeuPheSerGlySerGluGlnLeuLysProGlyAlaArgIle 146
Db 241 CCAGCTGGAGGCTCTGTCTCAGGCTCTGAGGCTCGCTGAAGCCCGGCGACCATC 300
QY 147 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 166
Db 301 TTCTCTTTGACGGCAAGGACGCTCTGAGGACCCACCTGGCCCGAGAGCGGTGG 360
QY 167 HisGlySerAspProSerGlyArgLeuThrAspSerTyrCysGluThrTrpArgThr 186
Db 361 CATGGCTCGGACCGGCGGAGGCTCTCTGCTGGGGGCGAGGCTCTGTGGGCGAG 480
QY 207 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 226
Db 481 AGTGGCGGAGCTGCCATCAGCGCTACATCGTCTCTGCATTGAGAACAGCTTCATGACT 540
QY 227 Ser 227
Db 541 GCC 543
RESULT 9
US-10-131-241-53
; Sequence 53, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 53
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-241-53
Alignment Scores:
Pred. No.: 3,37e-84 Length: 549
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 67.39% Indels: 0
Gaps: 0
US-09-938-391-2 (1-230) x US-10-131-241-53 (1-549)

QY 67 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 86
Db 61 GGCGCATCGCGGCGATCCGCGGCGGCGGACTTCCAGTGTCTCCAGCAGCGCGGCGGTG 120
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 121 GGCGTGGCGGCGACCTTCGCGGCTTCTCTGCGGCTGAGGACCTGTACAGCATC 180
QY 107 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 126
Db 181 GTGCGCGTGGCGACCGCGCGGCTGCTCAAGCTCAAGGACGAGCTGTGTTT 240
QY 127 ProSerTrpGluAlaLeuPheSerGlySerGluGlnLeuLysProGlyAlaArgIle 146
Db 241 CCAGCTGGAGGCTCTGTCTCAGGCTCTGAGGCTCGCTGAAGCCCGGCGACCATC 300
QY 147 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 166
Db 301 TTCTCTTTGACGGCAAGGACGCTCTGAGGACCCACCTGGCCCGAGAGCGGTGG 360
QY 167 HisGlySerAspProSerGlyArgLeuThrAspSerTyrCysGluThrTrpArgThr 186
Db 361 CATGGCTCGGACCGGCGGAGGCTCTCTGCTGGGGGCGAGGCTCTGTGGGCGAG 480
QY 207 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 226
Db 481 AGTGGCGGAGCTGCCATCAGCGCTACATCGTCTCTGCATTGAGAACAGCTTCATGACT 540
QY 227 Ser 227
Db 541 GCC 543
RESULT 10
US-09-873-676-30
; Sequence 30, Application US/09873676
; Patent No. US20020077289A1
; GENERAL INFORMATION:
; APPLICANT: Macdonald, Nicholas J.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use
; FILE REFERENCE: 05213-0378 (43170-259333)
; CURRENT APPLICATION NUMBER: US/09/873,676
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 30
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-676-30
Alignment Scores:
Pred. No.: 3,39e-84 Length: 552
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 67.39% Indels: 0
Gaps: 0
US-09-938-391-2 (1-230) x US-09-873-676-30 (1-552)

QY 47 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 66
Db 1 CACAGCCACCGGAGCTTCCAGCGGCTGTCCAGCTGTGGCTCAACAGCCCGCTGTCA 60

GENERAL INFORMATION:
; APPLICANT: Lo, Kin-Ming
; APPLICANT: Li, Yue
; APPLICANT: Gillies, Stephen D
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
; FILE REFERENCE: LEX-006C1
; CURRENT APPLICATION NUMBER: US/10/292,418
; PRIOR FILING DATE: 2002-11-12
; PRIOR FILING DATE: 1999-08-25
; PRIOR FILING DATE: 09/383,315
; PRIOR FILING DATE: 1998-08-25
; PRIOR FILING DATE: 06/097,883
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(552)
; OTHER INFORMATION: Endostatin
; US-10-292-418-34

Alignment Scores:
Pred. No.: 4,88e-98 Length: 552
Score: 959.00 Matches: 183
Percent Similarity: 99.46% Conservative: 0
Best Local Similarity: 99.46% Mismatches: 1
Query Match: 77.40% Indels: 0
DB: 12 Gaps: 0

US-09-938-391-2 (1-230) x US-10-292-418-34 (1-552)

QY 47 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAenSerProGlnPro 66
Db 1 CACACCACAGGACTTCCAGCGGTGCTGCACCTGGTGCCCTGAACAGCCGCGAGCG 60
QY 67 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAla 86
Db 61 GCGGCGATGCGAGCAATCCCGGGAGCGGACTTCAGTGTCTCCAGAGCGCGCGCG 120
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 121 GGGCTGGCCGCGACCTTCCGGGCTTCTGCTGTGGGGCTGAGGACCTCTACAGCATC 180
QY 107 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 126
Db 181 GTGCGCGCGCCGACCGCACCGGGGTGCGTCTCAACCTCAGGAGCAGGTTCTTTC 240
QY 127 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 146
Db 241 CCCAGCTGGAGGCTTATTCTCGGGCTCCAGGGCCAGCTGAAGCCCGGGGCGCGCATC 300
QY 147 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 166
Db 301 TTCTCTTTCAGCGCAGAGATGCTCTGAGCACCCTGGCCCTGGCCGGAAGAGCGTGTGG 360
QY 167 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 186
Db 361 CACGGCTCCGACCCAGCGCGCGCTGCTGACGACAGCTACTGCGAGAGCTGCGCGGACG 420
QY 187 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGln 206
Db 421 GAGGCCCCGGCGGACCGGGCAGGCTGCTGCTGCTGGCGGCGAGGCTGCTGGAGCAG 480
QY 207 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAenSerValMetThr 226
Db 481 GAGGCGCGAGCTGCCGCCACGCTTCTGTGTGTCTGTGCATCGAGAACAGCGTCATGACC 540
QY 227 SerPheSerLys 230
|||||

Db 541 TCCTTCTCAAG 552

RESULT 4
US-10-131-241-50
; Sequence 50, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Canine sp.
; US-10-131-241-50

Alignment Scores:
Pred. No.: 4,88e-98 Length: 552
Score: 959.00 Matches: 183
Percent Similarity: 99.46% Conservative: 0
Best Local Similarity: 99.46% Mismatches: 1
Query Match: 77.40% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-2 (1-230) x US-10-131-241-50 (1-552)

QY 47 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAenSerProGlnPro 66
Db 1 CACACCACAGGACTTCCAGCGGTGCTGCACCTGGTGCCCTGAACAGCCGCGAGCG 60
QY 67 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAla 86
Db 61 GCGGCGATGCGAGGACTCCCGGGAGCGGACTTCCAGTGTCTCCAGAGCGCGCGCG 120
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 121 GGGCTGGCGCGACCTTCCGGGCTTCTGCTGTGGGGCTGAGGACCTCTACAGCATC 180
QY 107 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 126
Db 181 GTGCGCGCGCGACCGCACCGGGGTGCGTCTCAACCTCAGGAGCAGGTTCTTTC 240
QY 127 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 146
Db 241 CCCAGCTGGAGGCTTATTCTCGGGCTCCAGGGCCAGCTGAAGCCCGGGGCGCGCATC 300
QY 147 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 166
Db 301 TTCTCTTTCAGCGCAGAGATGCTCTGAGCACCCTGGCCCTGGCCGGAAGAGCGTGTGG 360
QY 167 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 186
Db 361 CACGGCTCCGACCCAGCGCGCGCTGCTGACGACAGCTACTGCGAGAGCTGCGCGGACG 420
QY 187 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGln 206
Db 421 GAGGCCCCGGCGGACCGGGCAGGCTGCTGCTGCTGGCGGCGAGGCTGCTGGAGCAG 480
QY 207 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAenSerValMetThr 226
Db 481 GAGGCGCGAGCTGCCGCCACGCTTCTGTGTGTCTGTGCATCGAGAACAGCGTCATGACC 540

```
Pred. No.: 1,266-103 Length: 4551
Score: 1019.00 Matches: 187
Percent Similarity: 89.43% Conservative: 16
Best Local Similarity: 82.38% Mismatches: 20
Query Match: 82.24% Indels: 4
DB: 14 Gaps: 1

US-09-938-391-2 (1-230) x US-10-060-036-144 (1-4551)

Qy 1 ProTpaArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
Db 3874 CCTGGCGGCAGATGACATCTGGCGCAGCCCTCGCTGCGCCAGCCGACCCCTAC 3933
Qy 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
Db 3934 CCGGAGCCCGCACACACAGCTCTTACGTGACATCTGGCGCGCGGAGCCACACAGCCCA 3993
Qy 41 ProValHisThrHisThrHisHisGlnAspPheGlnLeuValLeuHisValAla 60
Db 3994 CCC-----GCCACAGCCACCGGACTTCCAGCGGCTGCTCCACCTGGTTGCG 4041
Qy 61 LeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPhe 80
Db 4042 CTCAACAGCCCTGTCTCAGCGGCATGCGGGGCATCGCGGGCGGCGGACTTCCAGCTTC 4101
Qy 81 GlnAlaAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100
Db 4102 CAGCAGCGCGGCGCTGGGGCTGGCGGGACCTTCCGCGCTTCTCTGCTCGCGCTG 4161
Qy 101 GlnAspLeuTyrSerIleValArgAlaAspArgThrGlyValProValValAsnLeu 120
Db 4162 CAGGACCTGTACAGCTGTGCGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4221
Qy 121 ArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlnLeu 140
Db 4222 AAGGACAGCTGTGTTTCCAGCTGGGAGGCTGTGTTCTCAGGCTCTGAGGGTCCGCTG 4281
Qy 141 LysProGlyAlaArgIlePheSerPheAspGlyValLeuGlnHisProAlaTrp 160
Db 4282 AAGCCCGGGGCGACGATCTTCTCTTACGGGAGGAGGCTCTGAGGAGGAGGAGGAG 4341
Qy 161 ProArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTyr 180
Db 4342 CCCAGAGAGCGTGTGGCATGCTCGGACCCCAAGCGGCGGCGGCGGCGGCGGCGGCGG 4401
Qy 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla 200
Db 4402 TGTGAGAGCTGGCGGAGGAGGCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4461
Qy 201 GlyArgLeuLeuGluGlnAlaAlaSerCysArgHisAlaPheValValLeuCysIle 220
Db 4462 GGCAGGCTCTGGGCGAGAGTGGCGGAGGCTGCCATCAGCCCTACATCTGCTCTGCATT 4521
Qy 221 GluAsnSerValMetThrSer 227
Db 4522 GAGAACAGCTTCATGACTGCC 4542

RESULT 2
US-09-880-107-2178
; Sequence 2178, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054

; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatenIn Ver. 2.1
; SEQ ID NO 2178
; LENGTH: 3394
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L22548
US-09-880-107-2178

Alignment Scores:
Pred. No.: 5,416-102 Length: 3394
Score: 1003.00 Matches: 185
Percent Similarity: 88.11% Conservative: 15
Best Local Similarity: 81.50% Mismatches: 23
Query Match: 80.95% Indels: 4
DB: 10 Gaps: 1

US-09-938-391-2 (1-230) x US-09-880-107-2178 (1-3394)

Qy 1 ProTpaArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
Db 1378 CCTGGCGGCAGATGACATCTGGCGCAGCCCTCGCTGCGCCAGCCGACCCCTAC 1437
Qy 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
Db 1438 CCGGAGCCCGCACACACAGCTCTTACGTGACATCTGGCGCGGCGGCGGCGGCGGCGG 1497
Qy 41 ProValHisThrHisThrHisHisGlnAspPheGlnLeuValLeuHisValAla 60
Db 1498 CCC-----GCCACAGCCACCGGACTTCCAGCGGCTGCTCCACCTGGTTGCG 1545
Qy 61 LeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPhe 80
Db 1546 CTCAACAGCCCTGTCTCAGCGGCATGCGGGGCATCGCGGGCGGCGGCGGCGGCGGCGG 1605
Qy 81 GlnAlaAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100
Db 1606 CAGCAGCGCGGCGCTGGGGCTGGCGGGACCTTCCGCGCTTCTCTGCTCGCGCTG 1665
Qy 101 GlnAspLeuTyrSerIleValArgAlaAspArgThrGlyValProValValAsnLeu 120
Db 1666 CAGGACCTGTACAGCTGTGCGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1725
Qy 121 ArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlnLeu 140
Db 1726 AAGGACAGCTGTGTTTCCAGCTGGGAGGCTGTGTTCTCAGGCTCTGAGGGTCCGCTG 1785
Qy 141 LysProGlyAlaArgIlePheSerPheAspGlyValLeuGlnHisProAlaTrp 160
Db 1786 AAGCCCGGGGCGACGATCTTCTCTTACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1845
Qy 161 ProArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTyr 180
Db 1846 CCCAGAGAGCGTGTGGCATGCTCGGACCCCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1905
Qy 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla 200
Db 1906 TGTGAGAGCTGGCGGAGGAGGCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1965
Qy 201 GlyArgLeuLeuGluGlnAlaAlaSerCysArgHisAlaPheValValLeuCysIle 220
Db 1966 GGCAGGCTCTGGGCGAGAGTGGCGGAGGCTGCCATCAGCCCTACATCTGCTCTGCATT 2025
Qy 221 GluAsnSerValMetThrSer 227
Db 2026 GAGAACAGCTTCATGACTGCC 2046

RESULT 3
US-10-292-418-34
; Sequence 34, Application US/10292418
; Publication No. US20030139365A1
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 17, 2003, 21:56:45 ; Search time 269.444 Seconds
(without alignments)
1910.324 Million cell updates/sec

Title: US-09-938-391-2
Perfect score: 1239
Sequence: 1 PWRADDILAGPRLDPPQY.....CRHAFVVLCIENSVMTSFSK 230

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1504479 seqs, 1118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame.p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09938391/runat_04082003_130658_27206/app_query.fasta_1.718
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09938391 -CCGN_1_129=runat_04082003_130658_27206
-NCPu=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1019	82.2	4551	14	US-10-060-036-144 Sequence 144, App

2	1003	81.0	3394	10	US-09-880-107-2178
3	959	77.4	552	12	US-10-292-418-34
4	959	77.4	552	14	US-10-131-241-50
5	891	71.9	632	14	US-10-131-241-51
6	843	68.0	573	10	US-09-998-831-12
7	835	67.4	546	14	US-10-042-347-4
8	835	67.4	549	12	US-10-292-418-3
9	835	67.4	549	14	US-10-131-241-53
10	835	67.4	552	9	US-09-873-676-30
11	831	67.1	540	14	US-10-131-241-48
12	831	67.1	558	9	US-09-775-174-3
13	831	67.1	558	9	US-09-775-325-3
14	831	67.1	565	13	US-10-036-869-37
15	828	66.8	552	12	US-10-292-418-17
16	828	66.8	624	13	US-10-080-797-4
17	827	66.7	551	13	US-10-080-797-2
18	817	65.9	534	14	US-10-042-347-6
19	817	65.9	537	14	US-10-131-241-59
20	634	51.2	900	14	US-10-131-241-45
21	542	43.7	574	14	US-10-060-036-64
22	528.5	42.7	5204	12	US-09-971-392-159
23	322	26.0	434	14	US-10-060-036-4
24	112.5	9.1	771	14	US-10-156-761-5483
c	112.5	9.1	9025608	14	US-10-156-761-1
26	110	8.9	287	14	US-10-060-036-97
27	105	8.5	1300	10	US-09-731-872-111
28	105	8.5	5065	14	US-10-067-457-4
29	104	8.4	1301	10	US-09-731-872-110
30	104	8.4	1515	14	US-10-156-761-7394
31	104	8.4	9025608	14	US-10-156-761-1
c	101.5	8.2	1344	9	US-09-815-242-7650
c	101.5	8.2	4522	14	US-10-171-581-163
c	101.5	8.2	1596	14	US-10-156-761-4136
c	100.5	8.1	1107	14	US-10-156-761-6709
36	100.5	8.1	2217	14	US-10-156-761-5547
37	100.5	8.1	2268	14	US-10-156-761-4417
38	100.5	8.1	3624	13	US-10-108-605-216
39	100	8.1	611	13	US-10-027-632-243061
40	100	8.1	611	13	US-10-027-632-243062
c	99.5	8.0	804	13	US-10-027-632-140701
c	99.5	8.0	804	13	US-10-027-632-150126
43	99	8.0	735	14	US-10-156-761-2117
c	99	8.0	26729	14	US-10-283-247-6
c	99	8.0	70383	14	US-10-283-247-3

ALIGNMENTS

RESULT 1

US-10-060-036-144
; Sequence 144, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 144
; LENGTH: 4551
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-036-144

Alignment Scores:

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
US-09-315-372-9

Alignment Scores:
Pred. No.: 0.682 Length: 3364
Score: 107.50 Matches: 57
Percent Similarity: 35.71% Conservative: 23
Best Local Similarity: 25.45% Mismatches: 48
Query Match: 8.68% Indels: 97
DB: 3 Gaps: 12

US-09-938-391-2 (1-230) x US-09-315-372-9 (1-3364)

```
Qy      8 LeuAlaGlyProProArg-----LeuLeuAspProGlnProTyrProGlyAlaProHis 25
Db      2187 CTGCCCCCCCCCCCCAACACACCTGCTGCTGAGCCTCACCCCCACCCCGGTGCTGGGTC 2246
Qy      26 HisGlySerTyrValHisPheGlnProAlaArgProThrGlyGlyProValHisThrHis 45
Db      2247 TTAGGCTCTGTACACCATGGAGAGAGCTCGC-----
Qy      46 ThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGln 65
Db      2280 -----TCTAAAAATAACCTGTC 2297
Qy      66 ProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAla 85
Db      2298 CCTGGTGGATCCAGG---GTAGGGGC-----AGGCTGAGGGCG 2333
Qy      86 AlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeu----- 103
Db      2334 GCC-----ACTTCCCTCAGCCGAGGTTGTTTCCCAAGAATGTTTCTGCTCTG 2387
Qy      104 -----TyrSerIleValArgAlaAspArgThrGlyValProValValAsnLeu 120
Db      2388 TAGCTTTTCTGTCAATCTG-----
Qy      121 ArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeu 140
Db      2409 -----CCATGGTGAGCAGCCTGC--ACTGGGCTTCTGGGAGAAACC 2448
Qy      141 LysProGlyAlaArgIle-----PheSerPheAspGlyArg----- 152
Db      2449 AAACCGGGTTCTAACCTTTTCAGTTACAGTTATTGCTTTCCTCTAGATGGGCGACTACAG 2508
Qy      153 -----AspValLeuGlnHisProAlaTrp----- 160
Db      2509 CCCCACCCACCCCGCTCTCTGTATCTCTCTGGGCTGGGATCCTAGGCTTTCAC 2568
Qy      161 -----ProArgLysSerVal----- 165
Db      2569 GGAAATTTCCCCCAGGTGTGTAGGTACGGCTCCCAAGAACAGTGTTCCT 2628
Qy      166 ---TrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrp 184
Db      2629 GGCAATGATGTTCTGAACCTCCAACCTGCAAAA---AATGACACATACCTTGACCCCTTG 2685
Qy      185 ArgThrGluAla 188
Db      2686 AAGCTGAGGCA 2697
```

Search completed: August 17, 2003, 23:00:06
Job time : 79.222 secs

87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIle 106
Db 124 GGGCTGTGGGACCTTCCTGGCTCTCTAGGCTGCAGGATCTCTATAGCATC 183
QY 107 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 126
Db 184 GTGGCGCTGTCTACCGGGGGTCTGTGCCCATCGTCAACCTGAAGACGAGGTGCTATCT 243
QY 127 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyAlaArgIle 146
Db 244 CCAGCTGGGACTCCCTGTTTCTGGCTCCAGGGTCAACTGCAACCCGGGGCCCGCATC 303
QY 147 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 166
Db 304 TTTTCTTTGACGGCAGAGATGCTCTGAGACACCCAGCCTGGCGGAGAGACGTATGG 363
QY 167 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 186
Db 364 CACGGCTCGGACCCAGTGGGGGAGGCTGATGGAGAGTTACTGTGAGACATGGCGAACT 423
QY 187 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 206
Db 424 GAAACTACTGGGGCTACAGTCAAGGCTCTCTCTGCTGTCAGCGAGGCTCTGGAAACAG 483
QY 207 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 226
Db 484 AAAGCTGGAGCTGCCACACACTACATGCTGCTGTGCATTTAGATAGCTTCATGACC 543
QY 227 SerPheSerLys 230
Db 544 TCTTTCTCCAAA 555
RESULT 11
US-08-985-526-37
; Sequence 37, Application US/08985526
; Patent No. 6080728
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA
; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Connolly, Bove, Lodge, & Hurz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,526
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorrow Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 565 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-985-526-37
Alignment Scores:
Pred. No.: 1,58e-72 Length: 565
Score: 831.00 Matches: 154
Percent Similarity: 92.93% Conservative: 17
Best Local Similarity: 83.70% Mismatches: 13
Query Match: 67.07% Indels: 0
DB: 3 Gaps: 0
US-09-938-391-2 (1-230) x US-08-985-526-37 (1-565)
QY 47 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 66
Db 11 CATACTCATCAGACTTTTCAGCCAGTGTCTCCACCTGGTGGCACTGAACCCCTGTCT 70
QY 67 GlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaAla 86
Db 71 GGAGGCATCGTGGTATCCGTGGAGCAGATTTCAGTGTCTCCAGCAAGCCGAGCCGTG 130
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 131 GGGCTGTGGGACCTTCCGGCTTTCCTCTCTAGGCTGCAGGATCTCTATAGCATC 190
QY 107 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 126
Db 191 GTGGCGCTGTGACCGGGGGTCTGTGCCCATCTCAACCTGAAGACGAGGTGCTATCT 250
QY 127 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyAlaArgIle 146
Db 251 CCCAGCTGGGACTCCCTGTTTCTGGCTCCAGGGTCAACTGCAACCCGGGGCCCGCATC 310
QY 147 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 166
Db 311 TTTTCTTTTGGCGGAGAGATGCTCTGAGACACCCAGCCTGGCGGAGAGCGGTATGG 370
QY 167 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 186
Db 371 CACGGCTCGGACCCCGAGTGGCGGAGGCTGATGGAGAGTTACTGTGAGACATGGCGAAT 430
QY 187 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 206
Db 431 GAAACTACTGGGGCTACAGGTGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 490
QY 207 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 226
Db 491 AAAGCTGGAGCTGCCACACAGCTACATGCTCTGTGCTGTGCTGTGCTGTGCTGTGCT 550
QY 227 SerPheSerLys 230
Db 551 TCTTTCTCCAAA 562
RESULT 12
US-09-315-689-6
; Sequence 6, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Folkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 6
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-315-689-6
Alignment Scores:
Pred. No.: 3.41e-71 Length: 534


```
QY 127 ProSerTripGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyAlaArgIle 146
Db 241 CCAGCTGGAGGCTCTGTCTCAGGCTCTGAGGTCCTGAGCCCGGGGCGACGCATC 300
QY 147 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTripProArgLysSerValTrip 166
Db 301 TTCTCTTTGACGCGCAAGACGCTCTGTAGGACACCCACCTGGCCCGAAGAGCGTGTGG 360
QY 167 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTripArgThr 186
Db 361 CATGGCTCGGACCCCAACGCGCGCAGGCTGACCGAGAGCTACTGTGAGACGCTGGCGGAGC 420
QY 187 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 206
Db 421 GAGGCTCCCTCGGCCACCGGCCAGGCTCTCTGCTGTGGGGGCGAGGCTCTCTGGGGCAG 480
QY 207 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 226
Db 481 AGTGCGCGAGCTGCCATCAGCGCTACATCGTCTCTGCTTGGAGAACAGCTTCATGACT 540
QY 227 Ser 227
Db 541 GCC 543

RESULT 9
US-09-449-293-3
; Sequence 3, Application US/09449293
; Patent No. 6267954
; GENERAL INFORMATION:
; APPLICANT: Abitol, Marc
; APPLICANT: Uteza, Yves
; APPLICANT: Menasche, Maurice
; APPLICANT: Bossard, Carine
; APPLICANT: Van Den Bergh, Loic
; APPLICANT: Bonnel, Sebastien
; APPLICANT: Prats, Herve
; APPLICANT: Honiger, Jiri
; APPLICANT: Neuner-Jehle, Martin
; TITLE OF INVENTION: INTRAOCULAR TRANSPLANTATION OF ENCAPSULATED CELLS
; FILE REFERENCE: 8076.202US01
; CURRENT APPLICATION NUMBER: US/09/449, 293
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Rattus rattus
US-09-449-293-3

Alignment Scores:
Pred. No.: 1.56e-72 Length: 558
Score: 831.00 Matches: 154
Percent Similarity: 92.93% Conservative: 17
Best Local Similarity: 83.70% Mismatches: 13
Query Match: 67.07% Indels: 0
DB: 3 Gaps: 0

US-09-938-391-2 (1-230) x US-09-449-293-3 (1-558)
QY 47 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 66
Db 4 CATACTCATCAGGACTTTCAGCAGTGTCTCCACTGTGTGGCACTGAACACCCCTGTCT 63
QY 67 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAla 86
Db 64 GGAGGCATGCGTGTATCCGTCGAGCAGATTTCCAGTGTCTCCAGCAAGCCCGGCGTG 123
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 124 GGGCTGTGCGGACCTTCCGGGCTTTCTCTGCTTCTAGGCTGAGGATCTCTATAGCATC 183
```

```
QY 107 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 126
Db 184 GTGCGCGTGTCTGACCGGGGGTCTGTGCCATCGTCAACCTGAGGACGAGGTGTATCT 243
QY 127 ProSerTripGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyAlaArgIle 146
Db 244 CCAGCTGGAGCTCTGTCTTCTGGCTCCCAAGGTCACCTGCAACCCCGGGGCGGCAATC 303
QY 147 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTripProArgLysSerValTrip 166
Db 304 TTTTCTTTTGACGCGCAGAGATGCTCTGACACCCAGCTGCGCCGAGAGCGGTATGG 363
QY 167 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTripArgThr 186
Db 364 CACGGCTCGGACCCCAAGGCTGGGGCGAGGCTGTATGGAGAGTTACTGTGAGACATGGCGAACT 423
QY 187 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 206
Db 424 GAACTTACTGGGGCTACAGGTGAGGCTCTCTCTGCTGTCTGAGGAGGCTCTCTGGAACAG 483
QY 207 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 226
Db 484 AAAGCTGCGAGCTGCCACACAGCTACATCGTCTGTGCTTGCATTCAGATAGCTTCATGACC 543
QY 227 SerPheSerLys 230
Db 544 TCTTTCTCCAAA 555

RESULT 10
US-09-775-325-3
; Sequence 3, Application US/09775325
; Patent No. 6500449
; GENERAL INFORMATION:
; APPLICANT: Abitol, Marc
; APPLICANT: Uteza, Yves
; APPLICANT: Menasche, Maurice
; APPLICANT: Bossard, Carine
; APPLICANT: Van Den Bergh, Loic
; APPLICANT: Bonnel, Sebastien
; APPLICANT: Prats, Herve
; APPLICANT: Honiger, Jiri
; APPLICANT: Neuner-Jehle, Martin
; TITLE OF INVENTION: INTRAOCULAR TRANSPLANTATION OF ENCAPSULATED CELLS
; FILE REFERENCE: 8076.202US01
; CURRENT APPLICATION NUMBER: US/09/775,325
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: US 09/449,293
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Rattus rattus
US-09-775-325-3

Alignment Scores:
Pred. No.: 1.56e-72 Length: 558
Score: 831.00 Matches: 154
Percent Similarity: 92.93% Conservative: 17
Best Local Similarity: 83.70% Mismatches: 13
Query Match: 67.07% Indels: 0
DB: 4 Gaps: 0

US-09-938-391-2 (1-230) x US-09-775-325-3 (1-558)
QY 47 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 66
Db 4 CATACTCATCAGGACTTTCAGCAGTGTCTCCACTGTGTGGCACTGAACACCCCTGTCT 63
QY 67 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAla 86
Db 64 GGAGGCATGCGTGTATCCGTCGAGCAGATTTCCAGTGTCTCCAGCAAGCCCGGCGTG 123
```

QY 83 AlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAsp 102
Db 130 GCCCGAGCGTGGGGTGTTCGGCACCTTCGGGGCTTTCTGCTCTTAGCTGCAGGAT 189
QY 103 LeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAsp 122
Db 190 CTCTATAGCATCGTGGCGCGTGTGACCGGGGGTCTGTGCCCCATCGTCAACCTGAAGGAC 249
QY 123 GluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLysPro 142
Db 250 GAGTGCTATCTCCAGCTGGGACTCCCTGTTTCTGGCTCCAGGGTCACTGCAACCC 309
QY 143 GlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArg 162
Db 310 GGGGCGCGCATCTTTCTTTGACGGCAGAGATGTCTGTAGACACCCAGCGCTGGCGCAG 369
QY 163 LysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGlu 182
Db 370 AAGAGCGTATGGACCGCTCGGACCCCGAGTGGCGGAGGCTGATGAGAGATTACTGTGAG 429
QY 183 ThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArg 202
Db 430 ACATGGGAATGAAACTACTGGGGCTACAGGTACAGCTCCCTCCCTGCTGTCAGGAGG 489
QY 203 LeuLeuGlnGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsn 222
Db 490 CTCCTGGACAGAAAGCTGCGAGTGCACACAGCTACATCTGCTGTCATTGAGAT 549
QY 223 SerValMetThrSerPheSerLys 230
Db 550 AGCTTCATGACCTTTCTTCCAAA 573

RESULT 7

US-09-315-689-4

; Sequence 4, Application US/09315689

; Patent No. 6346510

; GENERAL INFORMATION:

; APPLICANT: Folkman, Judah

; APPLICANT: O'Reilly, Michael

; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions

; FILE REFERENCE: 05213-0229

; CURRENT APPLICATION NUMBER: US/09/315,689

; CURRENT FILING DATE: 1999-05-20

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 546

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-315-689-4

Alignment Scores:
Pred. No.: 6,16e-73 Length: 546
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 67.39% Indels: 0
DB: 4 Gaps: 0

US-09-938-391-2 (1-230) x US-09-315-689-4 (1-546)

QY 47 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 66
Db 1 CACAGCACCGCGACTTCCAGCGGGTGTCTCCACTGGTTCGCTCAACAGCCCCCTGTCA 60
QY 67 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaA 86
Db 61 GGCAGCATGCGGGCATCCCGGGGCGGACTTCCAGTGTCTTCAGCAGCGCGGGCGGTG 120
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 121 GGGCTGGCGGACCTTCCTGCTCGCGCTTCCTGCTCGCGCTTCAGGAGCTGTACAGCATC 180

QY 107 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 126
Db 181 GTGGCGCGTGGCGACCGCGCAGCGTGGCCATCGTCAACCTCAAGGACGAGCTGTGTTT 240
QY 127 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyAlaArgIle 146
Db 241 CCAGCTGGAGGCTCTGTCTCAGGCTCTGAGGGTCCGCTGAAGCCCCGGGCGCAGCATC 300
QY 147 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 166
Db 301 TTCTCTTTGACGGCAGGACGCTCTGTGAGCACCCTCCCTGGCCCCCAGAGAGCGGTGG 360
QY 167 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 186
Db 361 CATGGCTCGGACCCCAACGGCGCAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGACG 420
QY 187 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGln 206
Db 421 GAGGCTCCCTCGGCCACCGGCCAGGCTCTCTGCTGCTGGGGGCGCAGGCTCCTGGGGCAG 480
QY 207 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 226
Db 481 AGTGGCGGAGCTGCCATCACGCTACATCGTGTCTGCAATTGAGAACAGCTTCATGACT 540
QY 227 Ser 227
Db 541 GCC 543

RESULT 8

US-09-206-059-30

; Sequence 30, Application US/09206059

; Patent No. 6201104

; GENERAL INFORMATION:

; APPLICANT: Macdonald, Nicholas

; APPLICANT: Sim, Kim Lee

; TITLE OF INVENTION: Angiogenesis-Inhibiting Protein Binding Peptides and

; TITLE OF INVENTION: Proteins and Methods of Use

; FILE REFERENCE: 05213-0370

; CURRENT APPLICATION NUMBER: US/09/206,059

; CURRENT FILING DATE: 1998-12-04

; NUMBER OF SEQ ID NOS: 80

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 30

; LENGTH: 552

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-206-059-30

Alignment Scores:
Pred. No.: 6,25e-73 Length: 552
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 67.39% Indels: 0
DB: 3 Gaps: 0

US-09-938-391-2 (1-230) x US-09-206-059-30 (1-552)

QY 47 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 66
Db 1 CACAGCACCGCGACTTCCAGCGGGTGTCTCCACTGGTTCGCTCAACAGCCCCCTGTCA 60
QY 67 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaA 86
Db 61 GGCAGCATGCGGGCATCCCGGGGCGGACTTCCAGTGTCTTCAGCAGCGCGGGCGGTG 120
QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
Db 121 GGGCTGGCGGACCTTCCTGCTCGCGCTTCCTGCTCGCGCTTCAGGAGCTGTACAGCATC 180
QY 107 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 126
Db 181 GTGGCGCGTGGCGACCGCGCAGCGCTGGCCATCGTCAACCTCAAGGACGAGCTGTGTTT 240

QY 140 LeuLysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAla 159
Db 3592 GTCAACCCGGGGCCGCACTCTTTTTCGAGCAGATGCTCTGAGACACCCAGCC 3651
QY 160 TrpProArgLysSerValTrpHisGlySerAspProSerGlyArgLeuThrAspSer 179
Db 3652 TGCCCGCAGAGAGCGTATGCGACGGCTCGGACCCCGAGTGGGGAGGCTGTAGAGAGT 3711
QY 180 TyrCysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeu 199
Db 3712 TACTGTGAGCATGGCGAACTGAACTACTGGGGCTACAGTCAGGCCTCTCCCTGCTG 3771
QY 200 AlaGlyArgLeuLeuGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCys 219
Db 3772 TCAGGCAGGCTCTCGAACAGAAAGCTCGGAGCTGCCAACACAGCTACATCGTCTGTGC 3831
QY 220 IleGluAsnSerValMetThrSerPheSerLys 230
Db 3832 ATTGAGAATAGCTTCATGACCTCTTCTCCAAA 3864
RESULT 3
US-09-561-500-12
; Sequence 12, Application US/09561500
; Patent No. 6342219
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002500
; CURRENT APPLICATION NUMBER: US/09/561,500
; PRIOR FILING DATE: 2000-04-28
; PRIOR FILING DATE: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; NAME/KEY: CDS
; LOCATION: (1)..(573)
US-09-561-500-12
Alignment Scores:
Pred. No.: 1.09e-73 Length: 573
Score: 843.00 Matches: 156
Percent Similarity: 92.02% Conservative: 17
Best Local Similarity: 82.98% Mismatches: 15
Query Match: 68.04% Indels: 0
Gaps: 4
US-09-938-391-2 (1-230) x US-09-561-500-12 (1-573)
QY 43 HisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsn 62
Db 10 CATCACCATCACCATACTCATCAGGACTTTCAGCCAGTGTCTCCACTGGTGCACTGAAC 69
QY 63 SerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnGln 82
Db 70 ACCCCCTCTCTGGAGCATCGTGGTATCCGTGGAGCAGATTTCCAGTCTTCAGCAA 129
QY 83 AlaArgAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAsp 102
Db 130 GCCCGAGCCGTGGGGTGTCTGGGCACCTTCGGGCTTCTCTCTCTAGGCTGCAGGAT 189
QY 103 LeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAsp 122
Db 190 CTCTATAGCATCTGGCCCGCTGTGACCCGGGGTCTGTGCCCCATCGTCAACCTGAAGGAC 249
QY 123 GluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLysPro 142

Db 250 GAGGTCTATCTCCAGCTGGGACTCCCTGTTTTCGGTCCAGGGTCAACTGCAACCC 309
QY 143 GlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArg 162
Db 310 GGCGCCCGCATCTTTCTTTTTCAGCGCAGAGATGCTCTGAGACACCCAGCTGGCCGAG 369
QY 163 LysSerValTrpHisGlySerAspProSerGlyArgLeuThrAspSerTrpCysGlu 182
Db 370 AAGAGCGTATGCGACGGCTCGGACCCCACTGGCGGAGGCTGATGAGAGTACTGTGAG 429
QY 183 ThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArg 202
Db 430 ACATGCGCAACTGAAACTACTGGGGCTACAGGTACAGGCTCTCTCCCTGTGTGAGCAGG 489
QY 203 LeuLeuGluGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsn 222
Db 490 CTCCTGGAAACAAAGAGCTCGAGCTGCCAACACAGCTACATCGTCTGTGCAATTGAGAT 549
QY 223 SerValMetThrSerPheSerLys 230
Db 550 AGCTTCATGACCTCTTCTCCAAA 573
RESULT 4
US-09-561-108-12
; Sequence 12, Application US/09561108
; Patent No. 6342221
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002584
; CURRENT APPLICATION NUMBER: US/09/561,108
; CURRENT FILING DATE: 2000-04-28
; PRIOR FILING DATE: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; NAME/KEY: CDS
; LOCATION: (1)..(573)
US-09-561-108-12
Alignment Scores:
Pred. No.: 1.09e-73 Length: 573
Score: 843.00 Matches: 156
Percent Similarity: 92.02% Conservative: 17
Best Local Similarity: 82.98% Mismatches: 15
Query Match: 68.04% Indels: 0
Gaps: 4
US-09-938-391-2 (1-230) x US-09-561-108-12 (1-573)
QY 43 HisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsn 62
Db 10 CATCACCATCACCATACTCATCAGGACTTTCAGCCAGTGTCTCCACTGGTGCACTGAAC 69
QY 63 SerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnGln 82
Db 70 ACCCCCTCTCTGGAGCATCGTGGTATCCGTGGAGCAGATTTCCAGTCTTCAGCAA 129
QY 83 AlaArgAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAsp 102
Db 130 GCCCGAGCCGTGGGGTGTCTGGGCACCTTCGGGCTTCTCTCTCTAGGCTGCAGGAT 189
QY 103 LeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAsp 122

SEQUENCE CHARACTERISTICS:
LENGTH: 3394
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-159-784-4

Alignment Scores:
Pred. No.: 2 95e-88 Length: 3394
Score: 1003.00 Matches: 185
Percent Similarity: 88.11% Conservative: 15
Best Local Similarity: 81.50% Mismatches: 23
Query Match: 80.95% Indels: 4
DB: 1 Gaps: 1

US-09-938-391-2 (1-230) x US-08-159-784-4 (1-3394)

```
QY 1 ProTpaArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
DB 1378 CCTGGGGGCGACATGACATCTCTGGCCAGCCCTCTGGGCTGCCCGAGCCCTTAC 1437
QY 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
DB 1438 CCCGAGGCGCCGACACACAGCTCTCTAGTGCACCTGGCGCCCGCCACGACCCCAAGCCCA 1497
QY 41 ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuValAla 60
DB 1498 CCC-----GCCACACCCACCGGACCTTCAGCCGGTGTCTCCACCTGGTGTGG 1545
QY 61 LeuAsnSerProGlnProGlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPhe 80
DB 1546 CTCACAGACCCCTGTGTCAGCGCGCATCGCGGCATCGCGGGGCCGACCTTCAGTCTTC 1605
QY 81 GlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100
DB 1606 CACGAGCGCGCGCGCTGGGGCTGGCGGCACCTTCGCGCCCTCTCTGCTCCGCGCTG 1665
QY 101 GlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValAlaAsnLeu 120
DB 1666 CAGACTGTACAGATCTGTCGCGCTGGCCACCGCGCAGCCGTGCCATCTGTCACCTTC 1725
QY 121 ArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeu 140
DB 1726 AAGCAGCAGCTGTGTTTCCAGCTGGAGGCTCTGTTCTCAGGCTCTGAGGGTCCGCTG 1785
QY 141 LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTrp 160
DB 1786 AAGCCCGGGGCGACGATCTTCTCTTGTACGGAAGGAGCTCTGAGGACCCACCTGG 1845
QY 161 ProArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTyr 180
DB 1846 CCCAGAGAGCTGTGGCATGCTGGACCCCAACCGGGCGCAGGCTGACGAGAGCTAC 1905
QY 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla 200
DB 1906 TGTGAGCGTGGCGAGCGAGGCTCCCTCGCCACCGGGCCAGGCTCTCTGCTGCTGGG 1965
QY 201 GlyArgLeuLeuGlnGlnGluAlaAlaSerCysArgHisAlaPheValLeuCysIle 220
DB 1966 GGCAGGCTCTGGGGCAGAGTGGCGGAGTGCATGCGCAGCTACATCGTCTCTGCTG 2025
QY 221 GluAsnSerValMetThrSer 227
DB 2026 GAGACAGCTTCATGACTGCC 2046
```

RESULT 2

US-08-159-784-1
Sequence 1, Application US/08159784
Patent No. 5643783
GENERAL INFORMATION:
APPLICANT: Bjorn R. Olsen
TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF
NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:
ADDRESSEE: Fieh & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/159,784
FILING DATE: December 1, 1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: John F. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00246/170001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-159-784-1

Alignment Scores:

Pred. No.:	5.84e-86	Length:	4031
Score:	980.50	Matches:	184
Percent Similarity:	87.88%	Conservative:	19
Best Local Similarity:	79.65%	Mismatches:	23
Query Match:	79.14%	Indels:	5
DB:	1	Gaps:	2

US-09-938-391-2 (1-230) x US-08-159-784-1 (1-4031)

```
QY 1 ProTpaArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
DB 3184 CCTGGCGAGCAGATGACATCTCTGGCCAAACCCACCGCGCTGCCAGACCCGAGCCTTAC 3243
QY 21 ProGlyAlaPro---HisHisGlySerTyrValHisPheGlnProAlaArgProThrGly 39
DB 3244 CTGGAGTTCACATCACCACAGATTCCTATGTGCACCTGCCCGCCAGCCGCCACCC 3300
QY 40 GlyProValHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuVal 59
DB 3301 -----CTCTCCTGCTCATCTCATCAGGACTTTTCAGCCAGTGTCTCCACCTGGTG 3351
QY 60 AlaLeuAsnSerProGlnProGlyMetArgGlyIleArgGlyAlaAspPheGlnCys 79
DB 3352 GCACCTGAACACCCCTCTGTCGAGGACATGCGTGTGTATCGGTGGAGCAGATTTCAGTGC 3411
QY 80 PheGlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArg 99
DB 3412 TTCCAGAACCCCGAGCCGCTGGGGCTGTGGGCACCTTCGCGGCTTCTCTCTCTAGG 3471
QY 100 LeuGlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValAlaAsn 119
DB 3472 CTGACGAGTCTCTATAGCATCGTGGCCGCTGTGACCCGGGGTCTGTGCCATCGTCAAC 3531
QY 120 LeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGln 139
DB 3532 CTGAAGGACGAGTGTATCTCTCCAGCTGGGACTCCCTCTCTGCTGCCAGGGGTCAA 3591
```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 17, 2003, 19:49:34 ; Search time 72.2222 Seconds
(without alignments)
1405.636 Million cell updates/sec

Title: US-09-938-391-2

Perfect score: 1239

Sequence: 1 PWRADDILGPPRLDPPQY.....CRHAFVWLCIENSVMTSFSK 230

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DRV=slp
-Q=/cgn2_1/USPTO.spool_p/US0938391/runat_04082003_130656_27125/app_query.fasta_1.718
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0938391@cgn_1_133 @runat_04082003_130656_27125 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1003	81.0	3394	1	US-08-159-784-4
2	980.5	79.1	4031	1	US-08-159-784-1
3	843	68.0	573	4	US-09-561-500-12
4	843	68.0	573	4	US-09-561-108-12
5	843	68.0	573	4	US-09-561-526-12
6	843	68.0	573	4	US-09-561-499-12
7	835	67.4	546	4	US-09-315-689-4
8	835	67.4	552	3	US-09-206-059-30
9	831	67.1	558	3	US-09-449-293-3
10	831	67.1	558	3	US-09-775-325-3
11	831	67.1	565	3	US-08-985-526-37
12	817	65.9	534	4	US-09-315-689-6
					Sequence 4, Appli
					Sequence 1, Appli
					Sequence 12, Appl
					Sequence 12, Appl
					Sequence 12, Appl
					Sequence 4, Appli
					Sequence 30, Appl
					Sequence 3, Appli
					Sequence 37, Appl
					Sequence 6, Appli

13	107.5	8.7	3364	2	US-08-735-609-9	Sequence 9, Appli
14	107.5	8.7	3364	2	US-08-735-609-9	Sequence 9, Appli
15	107.5	8.7	3364	3	US-09-315-372-9	Sequence 9, Appli
16	107.5	8.7	3364	3	US-09-244-752-9	Sequence 9, Appli
17	107.5	8.7	3364	3	US-09-245-497-9	Sequence 9, Appli
18	107.5	8.7	3364	4	US-09-562-919-9	Sequence 9, Appli
C 19	101.5	8.2	2457	2	US-08-723-415B-12	Sequence 12, Appl
C 20	101.5	8.2	2457	3	US-09-189-627A-12	Sequence 12, Appl
C 21	101.5	8.2	2457	4	US-09-710-861-12	Sequence 12, Appl
C 22	101.5	8.2	2457	1	US-08-306-691B-18	Sequence 18, Appl
C 23	101.5	8.2	2517	1	US-08-385-142-2	Sequence 2, Appli
C 24	101.5	8.2	2517	2	US-08-481-814A-1	Sequence 1, Appli
C 25	101.5	8.2	2517	3	US-08-801-092-2	Sequence 2, Appli
C 26	101.5	8.2	2517	3	US-09-517-584A-3	Sequence 3, Appli
C 27	101.5	8.2	2517	4	US-09-315-113-2	Sequence 2, Appli
C 28	99	8.0	1275	4	US-09-252-991A-16364	Sequence 16364, A
C 29	99	8.0	1698	4	US-09-252-991A-16474	Sequence 16474, A
C 30	98.5	7.9	1110	4	US-09-252-991A-8962	Sequence 8962, Ap
C 31	98.5	7.9	1278	4	US-09-252-991A-9043	Sequence 9043, Ap
C 32	98.5	7.9	1968	4	US-09-252-991A-8743	Sequence 8743, Ap
C 33	98.5	7.9	2456	1	US-07-882-711-1	Sequence 1, Appli
C 34	98.5	7.9	2456	2	US-08-462-174-1	Sequence 1, Appli
C 35	98.5	7.9	2492	1	US-08-139-937-13	Sequence 13, Appl
C 36	98.5	7.9	2492	5	PCT-US93-11310-13	Sequence 13, Appl
C 37	97.5	7.9	936	4	US-09-252-991A-11432	Sequence 11432, A
C 38	97.5	7.9	1500	4	US-09-252-991A-11439	Sequence 11439, A
C 39	97.5	7.9	1665	4	US-09-252-991A-11397	Sequence 11397, A
C 40	97.5	7.9	2218	2	US-08-985-090-4	Sequence 4, Appli
C 41	97.5	7.9	2218	3	US-09-165-543-31	Sequence 31, Appl
C 42	97.5	7.9	3244	3	US-09-165-543-4	Sequence 4, Appli
C 43	97	7.8	1420	3	US-09-362-506-1	Sequence 1, Appli
C 44	96.5	7.8	1095	4	US-08-914-372C-5	Sequence 5, Appli
C 45	96.5	7.8	1215	4	US-09-252-991A-9259	Sequence 9259, Ap

ALIGNMENTS

RESULT 1

US-08-159-784-4
; Sequence 4, Application US/08159784
; Patent No. 5643783
; GENERAL INFORMATION:
; APPLICANT: Bjorn R. Olsen
; TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,784
; FILING DATE: December 1, 1993
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: John F. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00246/170001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:

```

Db      546 CAGGCTCCTGGGGCAGAGTGCCCGGAGCTGCCCATCAGGCTACATCGTGTATGCAATTGA 605
Qy      221 uAsnSerValMetThrSer 227
Db      606 GAACAGCTTTCATGACTGCC 624

```

Search completed: August 17, 2003, 20:06:31
Job time : 283.222 secs

Score: 846.50 Matches: 162
Percent Similarity: 86.96% Conservativity: 18
Best Local Similarity: 78.26% Mismatches: 23
Query Match: 68.32% Indels: 4
DB: 22 Gaps: 2

US-09-938-391-2 (1-230) x AAH79104 (1-641)

QY 25 HisHisGlySerTyrrValHisPheGlnProAlaArgProThrGly-----GlyProVal 42
DB 6 CACCATGGGGTACTGCTACACAGAGGAGCGTCTCAGTCTGGTCTTGCCTCTCTGTT 65
QY 43 HisThrHis-----ThrHisThrHisGlnAspPheGlnLeuValHisLeuValAlaLe 61
DB 66 TCCAAAGCATGGGAGGCACAGCCAGCGACTTCCAGCGCGTCTCCACCTGGTGGCT 125
QY 61 uAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheG 81
DB 126 CAACAGCCCCCTGCTCAGCGGCGCATGGGGGCATCCCGGGGCGGACTTCCAGTGTCTCA 185
QY 81 nGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuG 101
DB 186 GCAGGCGGCGCGCTGGGGCTGGCGGCGACCTTCCGCGCTTCTGCTCTGGCGCTGCA 245
QY 101 nAspLeuTyrrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeuAr 121
DB 246 GGACCTGTACAGCATGCTGGCGCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCAA 305
QY 121 gAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLy 141
DB 306 GGACGAGCTGCTCTTCCAGCGTGGGAGGCTCTGTTCTCAGGCTCTGAGGCTCCGCTGAA 365
QY 141 sProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpPr 161
DB 366 GCCCGGGGCGCGCATCTTCTTAAACGCAAGGAGCGTCTGAGGACCCCGCGCGCGC 425
QY 161 oArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTyrrCy 181
DB 426 CCAGAGAGCGGTGGCATGGCTCGACCCCAACGCGCGCGCGCGCGCGCGCGCGCGCTACTG 485
QY 181 sGluThrTrpArgThrGluAlaProAlaThrGlyGlnAlaSerSerLeuLeuAlaG 201
DB 486 TGAGACGTGGCGGACGAGGCTCCCTCGGCGCACGGGCGCGCGCGCGCGCGCGCGCGCG 545
QY 201 yArgLeuLeuGluGlnAlaAlaSerCysArgHisAlaPheValValLeuCysIleG 221
DB 546 CAGGCTCTGGGCGAGGTGGCGGAGTGCAGGCTGCATCAGGCTACATCGTCTATGCAATTGA 605
QY 221 uAsnSerValMetThrSer 227
DB 606 GAACAGCTTCATGACTGCC 624

RESULT 15

AAL55454
ID AAL55454 standard; DNA; 641 BP.

XX AAL55454;

AC AAL55454;

XX 22-MAY-2003 (first entry)

DE Specific tumour cell proliferation related nucleic acid, SEQ ID No 24.

XX Recombination virus; proliferating; tumour cell; anti-oncogene;

KW proliferation; telomerase promoter; therapy; tumour; db.

XX Unidentified.

OS WO200300640-A1.

XX 23-JAN-2003.

XX 12-JUL-2002; 2002WO-CN00493.

PR 12-JUL-2001; 2001CN-0126113.
XX (QIAN/) QIAN Q.
PA Qian Q, Wu M, Shan S;
XX WPI; 2002-464081/17.
DR A specific proliferation in a tumour cell which can express an anti-
XX oncogene with high efficiency and the use of it -
XX Example 6; Page 30; 56pp; Chinese.

XX The invention relates to a recombination virus proliferating in a tumour
CC cell, which can express an anti-oncogene with high efficiency. The
CC invention also relates to the method of its proliferation. A telomerase
CC promoter controlling the transcription of at least one necessary gene for
CC a recombination virus proliferating, can make the virus optionally
CC proliferate in a tumour cell, which has the activity of telomerase and
CC basically does not proliferate in a normal cell without the activity of a
CC telomerase. The recombination virus can be used in therapy of many kinds
CC of tumours. This polynucleotide sequence represents a DNA sequence
CC relating to the specific proliferation in a tumour cell of the invention.

XX Sequence 641 BP; 104 A; 220 C; 201 G; 116 T; 0 other;

Alignment Scores:
Pred. No.: 5,498-63 Length: 641
Score: 846.50 Matches: 162
Percent Similarity: 86.96% Conservativity: 18
Best Local Similarity: 78.26% Mismatches: 23
Query Match: 68.32% Indels: 4
DB: 22 Gaps: 2

US-09-938-391-2 (1-230) x AAL55454 (1-641)

QY 25 HisHisGlySerTyrrValHisPheGlnProAlaArgProThrGly-----GlyProVal 42
DB 6 CACCATGGGGTACTGCTACACAGAGGAGCGTCTCAGTCTGGTCTTGCCTCTCTGTT 65
QY 43 HisThrHis-----ThrHisThrHisGlnAspPheGlnLeuValHisLeuValAlaLe 61
DB 66 TCCAAAGCATGGGAGGCACAGCCAGCGACTTCCAGCGCGTCTCCACCTGGTGGCT 125
QY 61 uAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheG 81
DB 126 CAACAGCCCCCTGCTCAGCGGCGCATGGGGGCATCCCGGGGCGGACTTCCAGTGTCTCA 185
QY 81 nGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuG 101
DB 186 GCAGGCGGCGCGCTGGGGCTGGCGGCGACCTTCCGCGCTTCTGCTCTGGCGCTGCA 245
QY 101 nAspLeuTyrrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeuAr 121
DB 246 GGACCTGTACAGCATGCTGGCGCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCTCAA 305
QY 121 gAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLy 141
DB 306 GGACGAGCTGCTTCTTCCAGCTGGGAGGCTCTGTTCTCAGGCTCTGAGGCTCCGCTGAA 365
QY 141 sProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpPr 161
DB 366 GCCCGGGGCGCGCATCTTCTTAAACGCAAGGAGCGTCTGAGGACCCCGCGCGCGC 425
QY 161 oArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTyrrCy 181
DB 426 CCAGAGAGCGGTGGCATGGCTCGACCCCAACGCGCGCGCGCGCGCGCGCGCGCGCTACTG 485
QY 181 sGluThrTrpArgThrGluAlaProAlaThrGlyGlnAlaSerSerLeuLeuAlaG 201
DB 486 TGAGACGTGGCGGACGAGGCTCCCTCGGCGCACGGGCGCGCGCGCGCGCGCGCGCGCG 545
QY 201 yArgLeuLeuGluGlnAlaAlaSerCysArgHisAlaPheValValLeuCysIleG 221

XX PF 24-FEB-2000; 2000WO-US04798.
 XX PR 24-FEB-1999; 99US-0121341.
 XX PR 25-FEB-1999; 99US-0121633.
 XX PR 18-NOV-1999; 99US-0166176.
 XX PA (FORD-) FORD HEALTH SYSTEM HENRY.
 XX PI Dou D, Chopp M, Wang L, Mikkelsen T;
 XX DR WPI; 2000-572016/53.
 XX DR P-PSDB; AAB08407.
 XX PT Use of kringle protein and kringle derived from plasminogen and
 PT composition comprising kringle proteins for treating tumor and
 PT atherosclerosis, arthritis and retinopathy
 XX PS Disclosure; Page 151; 163pp; English.
 XX CC The present sequence encodes a human polypeptide which is a potent
 CC angiogenesis inhibitor, designated KED. The KED polypeptide, kringle
 CC proteins, or a kringle derived from human tissue plasminogen activator
 CC (tPA) protein are used to inhibit angiogenesis. Kringle protein, KED
 CC or tPA are useful for treating tumours, as well as atherosclerosis,
 CC arthritis, retinopathy and other similar diseases. KED is also useful
 CC for the treatment of diseases such as bronchial vascular congestion,
 CC inflammatory bowel disease, adult respiratory distress syndrome,
 CC Castleman's disease, psoriasis, hepatitis, aneurysm, renal disease
 CC and haemangioma.
 XX SQ Sequence 816 BP; 156 A; 259 C; 254 G; 147 T; 0 other;

Alignment Scores:
 Pred. No.: 4,09e-63 Length: 816
 Score: 849.50 Matches: 167
 Percent Similarity: 75.20% Conservative: 18
 Best Local Similarity: 67.89% Mismatches: 36
 Query Match: 68.56% Indels: 25
 DB: 21 Gaps: 4

US-09-938-391-2 (1-230) x AAA64013 (1-816)

QY 2 TrpArgAlaAspAspIleLeuAlaGlyProProArg-----LeuLeuAspProGlnPro 19
 DB 85 TGGGCTGCCAGGAG-----CCCATAGACACAGCATTTCACTCCAGAGACA 132
 QY 20 TyrProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGly 39
 DB 133 AATCCACGGCGGGCTCTGGAAAAAATACTGCCGT---AACCTGATGTTGATAGGT 189
 QY 40 GlyPro----- 41
 DB 190 GGTCCCTGGTGTACAGCAAAATCCAGAAACTTTACGACTACTGTGATGCCCTCAG 249
 QY 42 ValHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeu 61
 DB 250 TGTGGGCCCATATGCACAGCCACCGGACTTCAGCGGGTGTCCACCTGGTTGGGCTC 309
 QY 62 AsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGln 81
 DB 310 AACAGCCCCCTGTTCAGCGGGCATCGGGGCGATCCGGGGCGCGCACTTCCAGTGTCTCCAG 369
 QY 82 GlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGln 101
 DB 370 CAGCGCGGGCGGTGGGGCGGACCTTCCTCGCCCTTCCTGTCTCGCGCTCAG 429
 QY 102 AspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeuArg 121
 DB 430 GACCTGTACAGCATCGTGGCGCGTGGCGAGCCGCGCAGCCGTGCCATCGTCACTCAAG 489
 QY 122 AspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLys 141

DB 490 GACGAGCTGCTGTTTCCAGCTGGGAGGCTCTGTTCTCAGGCTCTGAGGGTCCGCTGAAG 549
 QY 142 ProGlyValaAraIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpPro 161
 DB 550 CCGGGGGACGATCTTCTCTTCGCGCAGGAGCTCTGAGGACCCACCTGGCCC 609
 QY 162 ArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCys 181
 DB 610 CAGAAGAGCGTGTGGCATGGCTCGGACCCCAACGGGCGCAGGCTGACCCAGAGCTACTGT 659
 QY 182 GluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGly 201
 DB 670 GAGACGTGGCGACGAGGCTCTCTCGGCCACGGGCCAGGCTCTCTGCTGCTGGGGGC 729
 QY 202 ArgLeuLeuGluGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGlu 221
 DB 730 AGGCTCTGGGCGAGAGTGGCGGAGCTGCCATCAGCCCTACATCGTCTGCTGATTGAG 789
 QY 222 AsnSerValMetThrSer 227
 DB 790 AACAGCTTCATGACTGCC 807

RESULT 14
 AAH79104
 ID AAH79104 standard; DNA; 641 BP.
 XX AC AAH79104;
 XX DT 20-NOV-2001 (first entry)
 XX DE Human endostatin encoding polynucleotide.
 XX KW Human; endostatin; angiostatin; virus; tumour vascular development;
 XX OS Homo sapiens.
 XX PN CN1298947-A.
 XX PD 13-JUN-2001.
 PF 01-DEC-2000; 2000CN-0127680.
 PR 01-DEC-2000; 2000CN-0127680.
 PA (QIAN/) QIAN Q.
 PI Qian Q, Che X, Ceng X;
 PR WPI; 2001-503384/56.
 PT Virus with specific reproduction in a tumour well and effective
 PT expression of tumour angiogenesis inhibitor and its construction method
 PS Disclosure; Page 16 (Disclosure); 52pp; Chinese.
 CC The invention relates to developing a virus with highly-effective
 CC expression of tumour vascular genesis inhibiting factor. The virus
 CC comprises a nucleotide sequence encoding the vascular genesis inhibiting
 CC factor inserted into a non-essential proliferation region of a virus gene
 CC group. The virus selectively proliferates in a tumour cell such that
 CC along with replication of virus, the nucleotide sequence copy number of
 CC the encoded vascular genesis inhibiting factor is increased. The
 CC expression of the vascular genesis inhibiting factor inhibits tumour
 CC vascular development and inhibits formation, growth and transfer of
 CC tumour. The present sequence is that of a human endostatin polynucleotide
 CC sequence, useful to the invention.
 SQ Sequence 641 BP; 104 A; 220 C; 201 G; 116 T; 0 other;

Alignment Scores:
 Pred. No.: 5.49e-63 Length: 641

ID AA251309 standard; DNA; 552 BP.
 AC AA251309;
 XX
 DT 06-JUN-2000 (first entry)
 XX
 DE Canine angiogenesis inhibitor, endostatin DNA.
 DE
 XX Canine; immunoglobulin Fc fragment; endostatin; immunofusin;
 KW angiogenesis; inhibitor; cytostatic; antirheumatoid; antiarthritic;
 KW antipsoriatic; antidiabetic; ophthalmological; immunosuppressant;
 KW vasotropic; vulnary; treatment; antiarteriosclerosis; tumour;
 KW metastasis; atherosclerosis; psoriasis; rheumatoid arthritis;
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
 KW wound granulation; keloid scar; gene therapy; ds.
 XX
 OS Canis familiaris.
 XX
 XX Key Location/Qualifiers
 FT 1..552
 FT CDS
 FT /tag= a
 FT /product= "Endostatin"
 FT /note= "Does not include stop codon"
 FT /partial
 XX
 PN WO200011033-A2.
 XX
 XX 02-MAR-2000.
 PD
 XX
 XX 25-AUG-1999; 99WO-US19329.
 PF
 XX 25-AUG-1998; 98US-0097883.
 PR
 XX (LEXI-) LEXINGEN PHARM CORP.
 PA
 XX
 XX Lo K, Li Y, Gillies SD;
 PI
 XX WPI: 2000-237616/20.
 DR P-PSDB; AA70265.
 XX
 XX Novel fusion protein of angiotensin or endostatin and an immunoglobulin
 PT FC region, useful for treating conditions mediated by angiogenesis,
 PT such as rheumatoid arthritis, tumors and macular degeneration -
 PS
 XX Example 8; Pages 58-59; 68pp; English.
 XX
 CC The patent discloses a DNA molecule encoding a fusion protein comprising
 CC a signal sequence, an immunoglobulin Fc region, and an angiogenesis
 CC inhibitor selected from angiotensin, endostatin, a plasminogen fragment
 CC having angiotensin activity, a collagen XVIII fragment having endostatin
 CC activity, or combinations of them. The fusion protein (immunofusin) is
 CC used to inhibit angiogenesis and to treat diseases or conditions mediated
 CC by angiogenesis. Conditions that may be treated include solid tumours,
 CC blood born tumours, tumour metastasis, benign tumours including
 CC haemangiomas, acoustic neuromas, neurofibromas, trachomas and pyrogenic
 CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases
 CC e.g. diabetic retinopathy, retinopathy of prematurity, macular
 CC degeneration, corneal graft rejection, neovascular glaucoma, retrolental
 CC fibroplasia, rubeosis and Ogier-Webber syndrome; myocardial angiogenesis,
 CC plaque neovascularisation, telangiectasia, haemophilic joints,
 CC angiofibroma, wound granulation, and excessive or abnormal stimulation of
 CC endothelial cells, intestinal cells, atherosclerosis, scleroderma and
 CC hypertrophic scars, i.e. keloid scars. The DNA constructs may be used
 CC in gene therapy. The present sequence is a DNA encoding canine
 CC endostatin used in the construction of immunofusin containing canine
 CC immunoglobulin Fc fragment.
 XX
 SQ Sequence 552 BP; 77 A; 204 C; 190 G; 81 T; 0 other;

Alignment Scores:
 Pred. No.: 1.31e-72 Length: 552
 Score: 959.00 Matches: 183

Percent Similarity: 99.46% Conservative: 0
 Best Local Similarity: 99.46% Mismatches: 1
 Query Match: 77.40% Indels: 0
 DB: 21 Gaps: 0
 US-09-938-391-2 (1-230) x AA251309 (1-552)
 QY 47 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 66
 DB 1 CACACCCACAGGACTTCCAGCCGGTGTGACCTTGGTGGCCCTGACAGCCGCGAGCG 60
 QY 67 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 86
 DB 61 GCGCGCATCGAGGATCCGGGAGCGGACTTCCAGTGTCTCCAGCAGCGCGCGCG 120
 QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
 DB 121 GGGCTGGCGGCACCTTCCGGGCTTCTGTCTCGCGCTCGAGACCTCTACAGCATC 180
 QY 107 ValArgAlaAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 126
 DB 181 GTGCGCGCGCGACCGCACCGGGGTGCCCTGCTCACTCAGGACGAGGTGCTTTC 240
 QY 127 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLeuLeuProGlyAlaArgIle 146
 DB 241 CCCAGCTGGAGGCTTATCTCGGGTCCGAGGCGCAGCTGAAGCCCGGGCGCGCATC 300
 QY 147 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArglySerValTrp 166
 DB 301 TTCTCTTTTCGACGGCAGAGATGTCTGCAGCACCCCGCTGCCCGGGAAGAGCGTGTGG 360
 QY 167 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGlnThrTrpArgThr 186
 DB 361 CAGGCTCGACCCCGCGGGCGCGCTGACCGACGCTACTGCGAGACGTGGCGGAGC 420
 QY 187 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGluGln 206
 DB 421 GAGGCCCGCGGCGCACCGGCGAGCGTGTCTGCTGGCGGCGAGGCTGTGGAGCAG 480
 QY 207 GluAlaAlaSerCysArgHisAlaPheValLeuCysIleGluAsnSerValMetThr 226
 DB 481 GAGGCCGCGAGTGTGCGCACGCTTCTGTGTGTCTGTGCATCGAGAACAGCGTCATGACC 540
 QY 227 SerPheSerLys 230
 DB 541 TCCTTCTCCAG 552
 RESULT 13
 AAA64013
 ID AAA64013 standard; DNA; 816 BP.
 AC AAA64013;
 XX
 DT 20-DEC-2000 (first entry)
 DE
 DE DNA encoding a human angiogenesis inhibitor designated KED.
 XX
 KW Angiogenesis inhibitor; KED; kringle protein; angiogenesis; tPA;
 KW tissue plasminogen activator; tumour; atherosclerosis; arthritis;
 KW retinopathy; bronchial vascular congestion; inflammatory bowel disease;
 KW adult respiratory distress syndrome; Castelman's disease; psoriasis;
 KW hepatitis; aneurysm; renal disease; haemangioma; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT 1..816
 FT CDS
 FT /tag= a
 FT /product= "angiogenesis inhibitor KED"
 XX
 PN WO200049871-A1.
 XX
 PD 31-AUG-2000..

CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.

XX SQ Sequence 3394 BP; 668 A; 1130 C; 1080 G; 516 T; 0 other;

Alignment Scores:
Pred. No.: 2.16e-75 Length: 3394
Score: 1003.00 Matches: 185
Percent Similarity: 88.11% Conservative: 15
Best Local Similarity: 81.50% Mismatches: 23
Query Match: 80.95% Indels: 4
DB: 24 Gaps: 1

US-09-938-391-2 (1-230) x ABN95680 (1-3394)

QY 1 ProTrpAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
Db 1378 CCTGGCGGCAGATGACATCTCTGGCCAGCCCTGGCTGGCCGAGCCAGCCCTAC 1437

QY 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
Db 1438 CCGGAGGCCCGCACCATCTCTAGTGCATCGCGGCCGCGCAGCCACCAAGCCCA 1497

QY 41 ProValHisThrHisThrHisThrHisHisAspPheGlnLeuValLeuValAla 60
Db 1498 CCC-----GCCACGCCACCGGACATTCAGCGCGGTCTCCACCTGGTGG 1545

QY 61 LeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPhe 80
Db 1546 CTCAACAGCCCTGTGTCAGCGGCATCGCGGCATCGCGGCCGCGACATTCAGTCTTC 1605

QY 81 GlnAlaAlaArgAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeu 100
Db 1606 CAGCAGCGCGCGCGCGTGGCGGCGTGGCGGCACCTTCGCGCCCTCTCTCGCGCCTG 1665

QY 101 GlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeu 120
Db 1666 CAGGACTGTACAGATCTGTCGCGCGTGGCCAGCGCGCAGCCGTGCCATCTGCACTTC 1725

QY 121 ArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeu 140
Db 1726 AAGGACGAGTGTGTTCAGCTGGAGGCTCTGTCTCAGGCTCTGAGGGTCCGCTG 1785

QY 141 LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTrp 160
Db 1786 AAGCCCGGGGCAGCATCTCTCTTTCAGCGCAAGGAGTCTGAGGCACCCACCTGG 1845

QY 161 ProArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTyr 180
Db 1846 CCCAGAGAGCGTGTGGCATGCTGGACCCCAAGCGGCGAGGCTGACGAGCTAC 1905

QY 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla 200
Db 1906 TGTGACGCTGGCGGAGCGAGGCTCTCTCGGCGCAGCGGCGAGCCCTCTCGCTGGGG 1965

QY 201 GlyArgLeuLeuGlnGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCysIle 220
Db 1966 GGCAGGCTCTCGGGGAGAGTGGCGGAGTGGCGGAGTGCATCAGCCCTACATCGTCTCAT 2025

QY 221 GluAsnSerValMetThrSer 227
Db 2026 GAGAACAGCTTCATGATGCC 2046

RESULT 10
AAT84485
ID AAT84485 standard; cDNA; 4031 BP.
XX AC AAT84485;
XX AC AAT84485;
XX DT 19-NOV-1997 (first entry)
XX

DE Mouse alpha-1 collagen cDNA.
XX Alpha-1 collagen; type XVIII collagen; cartilage degeneration; ss.
XX Mus musculus.
XX Key Location/Qualifiers
CDS 1..3867 /tag= a
XX US5643783-A.
XX 01-JUL-1997.
XX 01-DEC-1993; 93US-0159784.
XX 01-DEC-1993; 93US-0159784.
XX (HARD) HARVARD COLLEGE.
XX Oh SP, Olsen BR;
XX WPI; 1997-350247/32.
XX P-PSDB; AAW26328.
XX Nucleic acid encoding human alpha-1 collagen - for production of recombinant alpha-1 collagen, for use in the treatment of cartilage degeneration
XX Disclosure; Fig 2; 35pp; English.
XX A cDNA clone (AAT84485) codes for a mouse novel type alpha-1 (XVIII) collagen (AAW26328) that is expressed in multiple organs, especially the liver, lung and kidney. It was isolated by screening murine 15.5- and 17.5-day embryo libraries with probes based on murine type XII collagen and on human alpha-1 collagen cDNA. Isolated nucleic acids can be used to express recombinant alpha-1 collagen in transformed host cells. Claimed nucleic acid (see AAT84484) coding for human alpha-1 collagen (AAW26327) can be used to treat patients suffering from diseases associated with degradation of cartilage, and for supplementing collagen.
XX Sequence 4031 BP; 873 A; 1172 C; 1266 G; 720 T; 0 other;

Alignment Scores:
Pred. No.: 2.16e-73 Length: 4031
Score: 980.50 Matches: 184
Percent Similarity: 87.88% Conservative: 19
Best Local Similarity: 79.65% Mismatches: 23
Query Match: 79.14% Indels: 5
DB: 18 Gaps: 2

US-09-938-391-2 (1-230) x AAT84485 (1-4031)

QY 1 ProTrpArgAlaAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
Db 3184 CCTGGCGAGCAGATGACATCTCTGGCCACCCAGCCGCTGCCAGCCGAGCCCTAC 3243

QY 21 ProGlyAlaPro---HisHisGlySerTyrValHisPheGlnProAlaArgProThrGly 39
Db 3244 CCGGAGTTCACATCACCACAGTTCCTATGTGCACCTCGCGCCGCGCCGCCACC--- 3300

QY 40 GlyProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuHisVal 59
Db 3301 -----CTCTCACTTGTCTCATCTCATCAGGACTTTCAGCCAGTGTCTCACCTGGTG 3351

QY 60 AlaLeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCys 79
Db 3352 GCACCTGAACACCCCTCTCTGGAGGCATGCTGTATCCGTGGAGCAGATTTCCAGTGC 3411

QY 80 PheGlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerArg 99
Db 3412 TTCCAGCAAGCCGAGCGGTGGGGCTGTGGGCACTTCCGGGCTTCTCTCTCTAGG 3471

PN W09931616-A1.
 XX 24-JUN-1999.
 XX 16-DEC-1998; 98WO-US26783.
 XX 16-DEC-1997; 97US-0069727.
 XX (HARD) HARVARD COLLEGE.
 XX Hohenester E, Olsen BR, Sasaki T, Timpl R;
 XX WPI; 1999-395243/33.
 XX P-PSDB; AAV25113.
 XX Identifying mimetics of mammalian endostatin
 XX Disclosure; Fig 5A-C; 75pp; English.
 XX This invention describes a novel method for identifying mimetics of
 CC mammalian endostatin. The method comprises identifying a compound
 CC having atomic coordinates with non-trivial similarity to selected
 CC coordinates of atoms of a mammalian endostatin involves (a) providing
 CC a library of atomic coordinates of compounds in a library of candidate
 CC compounds, (b) comparing the library of atomic coordinates to the
 CC selected coordinates of a mammalian endostatin and (c) selecting from the
 CC library at least one candidate compound on the basis of selection
 CC criteria which include similarities between the atomic coordinates of the
 CC selected candidate compound and the atomic coordinates of the mammalian
 CC endostatin. The invention also describes the use of an anti-angiogenic
 CC fragment of endostatin comprising a domain selected from a heparin
 CC binding domain, a receptor binding domain, and exposed on alpha-helix A
 CC domain, and a carbohydrate recognition domain (CRD) domain. The methods
 CC can be used for designing and selecting endostatin mimics. The compounds
 CC identified can be used for treating undesired angiogenesis, e.g. tumours.
 CC This sequence encodes human alpha(XVIII) collagen which is used in the
 CC description of the method.
 XX SQ Sequence 3394 BP; 668 A; 1130 C; 1080 G; 516 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,16e-75 Length: 3394
 Score: 1003.00 Matches: 185
 Percent Similarity: 88.11% Conservative: 15
 Best Local Similarity: 81.50% Mismatches: 23
 Query Match: 80.95% Indels: 4
 DB: 20 Gaps: 1
 US-09-938-391-2 (1-230) x AAX78379 (1-3394)
 QY 1 ProTIPArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAupProGlnProTyr 20
 DB 1378 CCTGGCGGCGAGATGATCTCTGGCGAGCCCTTGGGCTGGCGGCGCCAGCCCTAC 1437
 QY 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
 DB 1438 CCGGAGGCGCGGCGGCTGGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1497
 QY 41 ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuValAla 60
 DB 1498 CCC-----GCCACAGCCACCGGAGCTTCCAGCGGCTGGCTCCACCTGTGTGCG 1545
 QY 61 LeuAsnSerProGlnProGlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPhe 80
 DB 1546 CTAAACAGCCCTGTGTGGCGGCGGATCGGGGCGGATCGGGGCGGCGGCGGCGGCGGCGG 1605
 QY 81 GlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100
 DB 1606 CAGCAGCGCGGCGGCGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1665
 QY 101 GlnAspLeuTyrSerIleValArgAlaAspArgThrGlyValProValValAsnLeu 120
 DB 1666 CAGGACCTGTACAGCATCGTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1725

121 ArgAspGluValLeuPheProSerTIPGluAlaLeuPheSerGlySerGluGlyGlnLeu 140
 DB 1726 AAGGACGAGCTGCTGTTCCAGCTGGGAGGCTGTTCTCAGGCTCTGAGGGTCCGCTG 1785
 QY 141 LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTyr 160
 DB 1786 AAGCCCGGGGACGCGATCTTCTCTGACGCAAGAGAGCTCTGAGGACCCACCTGG 1845
 QY 161 ProArgLysSerValTyrHisGlySerAspProSerGlyArgGluLeuThrAspSerTyr 180
 DB 1846 CCCAGAGAGCGTGTGGCATGGCTCGGACCCCAAGGCGGCGGCTGACCCGAGACTAC 1905
 QY 181 CysGluThrTyrArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla 200
 DB 1906 TGTGAGACGTGGCGGACGAGGCTCCCTCGGCGGCGGCGGCGGCTCTCTGCTGCTGGG 1965
 QY 201 GlyArgLeuLeuGluGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCysIle 220
 DB 1966 GGCAGGCTCTGGGCGGAGAGTCCGCGAGCTGCCATCAGCTACATCTGCTGCTGCTG 2025
 QY 221 GluAsnSerValMetThrSer 227
 DB 2026 GAGACAGCTTCATGACTGCC 2046
 RESULT 9
 ABN95680
 ID ABN95680 standard; DNA; 3394 BP.
 AC ABN95680;
 XX 13-AUG-2002 (first entry)
 DE Gene #2178 used to diagnose liver cancer.
 KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 KW metastatic liver tumour; cytostatic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 OS Homo sapiens.
 XX WO200229103-A2.
 PD 11-APR-2002.
 XX 02-OCT-2001; 2001WO-US30589.
 XX 02-OCT-2000; 2000US-237054P.
 (GENE-) GENE LOGIC INC.
 Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 WPI; 2002-426119/45.
 Diagnosing and detecting the progression of liver cancer,
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample -
 XX Claim 1; SEQ ID NO 2178; 298pp; English.
 The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,

QY 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla 200
 Db 3921 TGTGACGCTGGCGGACGAGGCTCCCTCGGCCACGGGCAGGCTCTCTCGTGTGGG 3980
 QY 201 GlyArgLeuLeuGluGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCysIle 220
 Db 3981 GGAGGCTCTGGGGCAGAGTGCCGCGAGTGCCATCAGCCCTACATCGTGTCTGCAAT 4040
 QY 221 GluAsnSerValMetThrSer 227
 Db 4041 GAGAACAGCTTCATGACTGCC 4061

RESULT 7

AAT84484

ID AAT84484 standard; cDNA; 3394 BP.

XX

AC AAT84484;

XX

DT 19-NOV-1997 (first entry)

XX

DE Human alpha-1 collagen cDNA.

XX

KW Alpha-1 collagen; type XVIII collagen; cartilage degeneration; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..2055

FT /*tag= a

FT

XX

PN US5643783-A.

XX

PD 01-JUL-1997.

XX

PF 01-DEC-1993; 93US-0159784.

XX

PR 01-DEC-1993; 93US-0159784.

XX

PA (HARD) HARVARD COLLEGE.

XX

PI Oh SP, Olsen BR;

XX

DR WPI; 1997-350247/32.

XX

DR P-PSDB; AAW26327.

XX

PT Nucleic acid encoding human alpha-1 collagen - for production of

XX

PT recombinant alpha-1 collagen, for use in the treatment of cartilage

XX

PT degeneration

XX

PS Claim 1; Column 23-30; 35pp; English.

XX

CC A cDNA clone (AAT84484) codes for a human novel type alpha-1 (XVIII)

XX

CC collagen (AAW26327) that is expressed in multiple tissues, especially

XX

CC liver, lung and kidney. It was isolated from a placental cDNA

XX

CC library using a probe based on an unidentified collagenous protein

XX

CC and a probe based on mouse alpha-1 cDNA clone mc19. A claimed

XX

CC plasmid comprising alpha-1 collagen nucleic acid and an expression

XX

CC control sequence can be used to express recombinant collagen in

XX

CC prokaryotic or eukaryotic (especially mammalian) host cells. The

XX

CC collagen may be used to treat a patient suffering from a disease

XX

CC associated with degradation of cartilage, and for supplementing

XX

CC collagen.

XX

SQ Sequence 3394 BP; 668 A; 1130 C; 1080 G; 516 T; 0 other;

US-09-938-391-2 (1-230) x AAT84484 (1-3394)

QY 1 ProTrpArgAlaAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20

Db 1378 CCTCGCGGGCAGATGACATCTCTGGCCAGCCCTCTGGGCTGCCGAGCCCGAGCTTAC 1437

QY 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40

Db 1438 CCGGAGGCGCGCACACAGCTCTTACGTGACTCGCGGCCGCGCAGCCACAGCCCA 1497

QY 41 ProValHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuValAla 60

Db 1498 CCC-----GCCACACAGCCACCGGACTTCCAGCCGCTGCTCCACCTGTTGG 1545

QY 61 LeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPhe 80

Db 1546 CTCACAGCCCTCTGTAGGCGGCATGCGGGGCATCCCGGGGCGGACCTTCAGTGTCTTC 1605

QY 81 GlnGlnAlaArgAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100

Db 1606 CAGCAGGCGGCGGCTGGGGCTGGGGGCACCTTCCGGCCCTTCTGTCTCTCGCGCTG 1665

QY 101 GlnAspLeuTyrSerIleValArgAlaAspArgThrGlyValProValValLeuLeu 120

Db 1666 CAGGACCTGTACAGCATGCTGCGCGCTGCGGACCGCGCAGCGCTGCCATCTCAACCTC 1725

QY 121 ArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeu 140

Db 1726 AAGGACGAGCTGCTGTTCCTCCAGCTGGGAGGCTCTGTCTCAGGCTCTGAGGCTCGCTG 1785

QY 141 LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTrp 160

Db 1786 AAGCCCGGGGCGACGCTCTCTCTTCAGCGCAGGAGCTCTCTGAGGACCCACCTGG 1845

QY 161 ProArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTyr 180

Db 1846 CCCCAAGAGAGCTGTGGCTGCGGACCGCCCAACGGCGCGAGGCTGACCGAGAGTAC 1905

QY 181 CysGluThrTrpArgThrGluAlaProAlaAlaSerCysArgHisAlaPheValLeuAla 200

Db 1906 TGTGAGCTGCGGCGGAGGCTCCCTCGGCCACGGCCAGGCTCTCTGCTGCTGGGG 1965

QY 201 GlyArgLeuLeuGlnGlnAlaAlaSerCysArgHisAlaPheValLeuCysIle 220

Db 1966 GGCAGGCTCTGGGCGCAGAGTGCCGCGAGCTGCCATCAGCCCTACATCGTCTCTGCAAT 2025

QY 221 GluAsnSerValMetThrSer 227

Db 2026 GAGAACAGCTTCATGACTGCC 2046

RESULT 8

AAX78379

ID AAX78379 standard; cDNA; 3394 BP.

XX

AC AAX78379;

XX

DT 25-AUG-1999 (first entry)

XX

DE Human alpha1 (XVIII) collagen cDNA.

XX

KW Alphas(XVIII) collagen; mimetic; endostatin; atomic coordinate; library;

XX

KW anti-angiogenic; heparin binding domain; receptor binding domain; mimic;

XX

KW alpha-helix A domain; carbohydrate recognition domain; CRD domain;

XX

KW treatment; angiogenesis; tumour; human; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..2055

FT /*tag= a

FT /product= "alpha1 (XVIII) collagen"

FT /note= "partial sequence, no start codon given"

FT

XX

QY 1 ProTProArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
 DB 2386 CCTGGCGGGAGATGATCTCTGGCCAGCCCTCTGGCTGCCGAGCCCGAGCCCTAC 2445

QY 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
 DB 2446 CCGGAGCCCGCCACACAGCTCTTACGTGACCTCGCGCGGCGCGACCCACAGCCCA 2505

QY 41 ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValHisValAla 60
 DB 2506 CCC-----GCCACAGCCACCCGCGCTTCCAGCCGCTGCTCCACTGGTGG 2553

QY 61 LeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPhe 80
 DB 2554 CTCACAGCCCTCTGTGAGCGGCATCGGGGCATCCGGGGGCCGACTTCAGTGTTC 2613

QY 81 GlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100
 DB 2614 CAGCAGCGCGGGCCGTGGGGCTGGCGGCACCTTCGGCGCTTCTGTCCTCGCGCTG 2673

QY 101 GlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeu 120
 DB 2674 CAGGACCTGTACAGCATCTCTCTTTGACGCGCGCTGCGACCGCGCGCTGCTCAACCTC 2733

QY 121 ArgAspGluValLeuPheProSerTyrProGluAlaLeuPheSerGlySerGluGlyGlnLeu 140
 DB 2734 AAGCAGAGCTGTCTTCCAGCTGGAGGCTCTGTTCTCAGGCTCTCAGGGTCCGCTG 2793

QY 141 LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTyr 160
 DB 2794 AAGCCCGGGGACGATCTCTCTTTGACGCGAGGCTCTCAGGAGGCTCTCAGGAGCCACCTGG 2853

QY 161 ProArgLysSerValTyrHisGlySerAspProSerGlyArgArgLeuThrAspSerTyr 180
 DB 2854 CCCACAGAGCGGTGTGGCATGCTCGGACCCACGCGGGCGAGGTGACCGAGAGCTAC 2913

QY 181 CysGluThrTyrArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla 200
 DB 2914 TGTGAGACGTGGCGGACGAGGCTCTCTCGGCCACGGCCGAGGCTCTCTGCTGGGG 2973

QY 201 GlyArgLeuLeuGluGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCysIle 220
 DB 2974 GGCAGGCTCTGGCGAGAGTGGCGGAGTGGCGAGTGGCATCACGCTTACATCGTGTCTGCAIT 3033

QY 221 GluAsnSerValMetThrSer 227
 DB 3034 GAGAACAGCTTCATGACTGCC 3054

RESULT 6
 AAX7720
 ID AAX77720 standard; DNA; 5408 BP.
 XX AC AAX77720;
 XX AC AAX77720;
 DT 10-AUG-1999 (first entry)
 XX DE Human collagen 18 DNA.
 XX KW plasminogen; human; angiotensin; endostatin; gene therapy; collagen 18;
 KW anti-angiogenic; attenuation; cyclostatic; anti-diabetic; ophthalmology;
 KW tumour growth; solid tumour; diabetic retinopathy; retina; vector; ss.
 XX OS Homo sapiens.
 XX PN W09926480-A1.
 XX PD 03-JUN-1999.
 XX PF 20-NOV-1998; 98WO-US24950.
 XX PR 20-NOV-1997; 97US-0975424.
 XX XX

(GENE-) GENETIX PHARM INC.
 (MASI) MASSACHUSETTS INST TECHNOLOGY.
 Bachelot T, Leboulch P, Pawliuk RJ;
 WPI; 1999-357696/30.
 P-PSDB; AAY08694.
 Anti-angiogenic gene therapy vectors
 Disclosure; Page 75-76; 83pp; English.
 This invention describes a novel viral gene therapy vector comprising a nucleic acid molecule encoding an anti-angiogenic polypeptide chosen from human or murine angiotensin, human or murine endostatin and angiogenesis-inhibiting fusions and fragments, where the viral vector is sufficiently attenuated for use in human gene therapy. The products of the invention have anti-angiogenic, cytostatic, anti-diabetic and ophthalmological activity. The vector is used in gene therapy for inhibiting tumour growth in humans harbouring a solid tumour. The vector expresses an anti-angiogenic polypeptide. An additional use comprises treatment of diabetic retinopathy, where the anti-angiogenic polypeptide inhibits angiogenesis in the vicinity of the retina. The vector is administered to cells ex vivo and then administered to the patient.

XX SQ Sequence 5408 BP; 986 A; 1820 C; 1796 G; 806 T; 0 other;

Alignment Scores:
 Pred. No.: 1,668-76 Length: 5408
 Score: 1019.00 Matches: 187
 Percent Similarity: 89.43% Conservative: 16
 Best Local Similarity: 82.38% Mismatches: 20
 Query Match: 82.24% Indels: 4
 DB: 20 Gaps: 1

US-09-938-391-2 (1-230) x AAX77720 (1-5408)

QY 1 ProTProArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
 DB 3393 CCTGGCGGGAGATGATCTCTGGCCAGCCCTCTGGCTGCCGAGCCCGAGCCCTAC 3452

QY 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
 DB 3453 CCGGAGCCCGCCACACAGCTCTTACGTGACCTCGCGCGGCGCGACCCACAGCCCA 3512

QY 41 ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValHisValAla 60
 DB 3513 CCC-----GCCACAGCCACCCGCGACTTCCAGCCGCTGCTCCACTGGTGG 3560

QY 61 LeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPhe 80
 DB 3561 CTCACAGCCCTCTGTGAGCGGCATGCGGGGCATCCGGGGGCCGACTTCCAGTGTTC 3620

QY 81 GlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100
 DB 3621 CAGCAGCGCGGGCCGTGGGGCTGCGGGGCACCTTCCGCGCTTCTGTCCTCGCGCTG 3680

QY 101 GlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeu 120
 DB 3681 CAGGACCTGTACAGCATCTGTGCGCCGTGCGACCGCGAGCGGTGCCATCTGTCACCTC 3740

QY 121 ArgAspGluValLeuPheProSerTyrGluAlaLeuPheSerGlySerGluGlyGlnLeu 140
 DB 3741 AAGGACGAGCTGTCTTCCAGCTGGAGGCTCTGTTCTCAGGCTCTCAGGGTCCGCTG 3800

QY 141 LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTyr 160
 DB 3801 AAGCCCGGGGACGATCTTCTCTTTGACGCGAGGAGCTCTCAGGAGCCACCTGG 3860

QY 161 ProArgLysSerValTyrHisGlySerAspProSerGlyArgArgLeuThrAspSerTyr 180
 DB 3861 CCCACAGAGCGGTGTGGCATGCTCGGACCCACGCGGGCGAGGTGACCGAGAGCTAC 3920

Alignment Scores:

Pred. No.: 1.35e-76 Length: 4551
Score: 1019.00 Matches: 187
Percent Similarity: 89.43% Conservative: 16
Best Local Similarity: 82.38% Mismatches: 20
Query Match: 82.24% Indels: 4
DB: 24 Gaps: 1

US-09-938-391-2 (1-230) x ABN85301 (1-4551)

QY 1 ProTirArgAlaAspSerPheLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
DB 3874 CCTGGCGGCAGATGACATCTCTGGCCAGCCCTCCCTCGCTGCCGCCAGCCCTAC 3933
QY 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
DB 3934 CCGGAGCCCCGACACACAGCTCTTAGTGACCTGGCGCGCGCGGACCCACACGCCCA 3993
QY 41 ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuAla 60
DB 3994 CCC-----GCCACAGCCAGCCGACCTCCAGCCGCTGCCAGCTGGTTCG 4041
QY 61 LeuAsnSerProGlnProGlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPhe 80
DB 4042 CTCACAGCCCTCTGTGAGCGGCATCGCGGCATCCGCGGCGGCGGACTTCCAGTCTTC 4101
QY 81 GlnGlnAlaArgAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100
DB 4102 CAGCAGCGCGCGCTGGGGCTGGCGGCACCTTCGCGCCCTCTCTGCTCGCCCTG 4161
QY 101 GlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeu 120
DB 4162 CAGGACCTGTACAGATCTGTGCGCTGCGCCAGCCGCGCAGCTGCCCTCAACCTC 4221
QY 121 ArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeu 140
DB 4222 AAGGACGAGCTGTGTTCACAGCTGGAGGCTGTGTCTCAGGCTCTGAGGCTCGCGT 4281
QY 141 LysProGlyValAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTrp 160
DB 4282 AAGCCCGGGCAGCATCTCTCTTGGAGGCAAGAGCTCTGAGGACCCACCTGG 4341
QY 161 ProArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTrp 180
DB 4342 CCCCAGAGAGCTGTGGCATGCTCGACCCCAACCGCGCAGGCTGACCGAGAGCTAC 4401
QY 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla 200
DB 4402 TGTGAGACGTGGCGAGCGAGGCTCTCGGCCACCGCGCGAGGCTCTCTGCTGGGG 4461
QY 201 GlyArgLeuLeuGluGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCysIle 220
DB 4462 GGAGAGCTCTGGGCGAGAGTGGCGGAGTGCATCATCGCTACATCGTCTGCTGCTG 4521
QY 221 GluAsnSerValMetThrSer 227
DB 4522 GAGACAGCTTCATGACTGCC 4542

RESULT 5

ABQ54955
ID ABQ54955 standard; cDNA; 4875 BP.

XX AC ABQ54955;

XX DT 22-AUG-2002 (first entry)

XX DE Human ovarian antigen HEBB29 cDNA, SEQ ID NO:835.

XX KW Human: ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
XX KW inflammatory condition; immune disorder; blood disorder;

cardiovascular disorder; respiratory disorder; neurological disorder;
gastrointestinal disorder; urinary system disorder; drug screening;
gene therapy; chromosome mapping; forensic analysis;
antibody preparation; cytostatic; immunomodulatory; neuroprotective;
antiinflammatory; gynaecological; reproductive; chromosome 21q22.3;
gene; ss.

OS Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US18569.

XX 07-JUN-2000; 2000US-209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI: 2002-147878/19.

XX P-PSDB; ABP41878.

Isolated nucleic acid molecules encoding novel ovarian polypeptides,
useful in the prevention, treatment and diagnosis of cancer (e.g.
ovarian cancer), immune disorders, cardiovascular disorders and
neurological diseases -

Claim 1; SEQ ID No 835; 2922pp; English.

The invention relates to 2175 novel human ovarian antigens (ABP41054-
ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
encompasses polypeptides 90% identical and polynucleotides 95% identical
to the sequences of the invention. The invention additionally relates to
recombinant vectors and host cells comprising human ovarian antigen
polynucleotides, antibodies against human ovarian antigens, and the use
of ovarian antigen polynucleotides and polypeptides in diagnosing,
treating, prognosing or preventing various ovary and/or breast-related
disorders. Such conditions include ovarian cancer and breast cancer, and
metastatic tumours of ovarian or breast origin, reproductive system
disorders (e.g., infertility, disorders of pregnancy, anovulation,
polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
vaginitis), immune disorders (e.g., congenital and acquired
immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
blood-related disorders (e.g., anaemia), cardiovascular disorders,
respiratory disorders, neurological disorders, gastrointestinal disorders
and urinary system disorders. Ovarian antigen polypeptides and
polynucleotides may also be used in screening for compounds which
modulate ovarian antigen expression or activity. The polynucleotides may
further be used for gene therapy, chromosome mapping, in the
identification of individuals and in forensic analysis, and the
polypeptides may be used as food additives or to prepare antibodies
useful in disease diagnosis, drug targeting and phenotyping. The present
sequence represents cDNA encoding a human ovarian antigen of the
invention.

Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 4875 BP; 892 A; 1569 C; 1570 G; 829 T; 15 other;

Alignment Scores:

Pred. No.: 1.47e-76 Length: 4875
Score: 1019.00 Matches: 187
Percent Similarity: 89.43% Conservative: 16
Best Local Similarity: 82.38% Mismatches: 20
Query Match: 82.24% Indels: 4
DB: 24 Gaps: 1

US-09-938-391-2 (1-230) x ABQ54955 (1-4875)


```

PR 12-JUL-2001; 2001US-305484P.
PR 20-AUG-2001; 2001US-313999P.
PR 27-NOV-2001; 2001US-333626P.
XX (CORI-) CORIXA CORP.
XX Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
XX WPI: 2002-627435/67.
DR P-PSDB; ABP68617.
XX New isolated polynucleotide and pancreatic tumor polypeptides, useful
PT for diagnosing, preventing and/or treating cancer, particularly
PT pancreatic cancer
XX Claim 1; SEQ ID NO 144; 300pp + Sequence Listing; English.
XX The invention relates to an isolated polynucleotide (I) comprising: (a)
CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145);
CC (b) complements of (a); (c) sequences consisting of at least 20
CC contiguous residues of (a); (d) sequences that hybridize to (a), under
CC moderately stringent conditions; (e) sequences having at least 75% or 90%
CC identity to (a); or (f) degenerate variants of (a). Polypeptides
CC (ABP68596-ABP68637) encoded by (I) and oligonucleotide can be used to
CC detect cancer in a patient and compositions comprising polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations and
CC antigen presenting cells expressing the polypeptide are useful in
CC treating pancreatic cancer and stimulating an immune response. The
CC polynucleotides can be used as probes or primers for nucleic acid
CC hybridization, in the design and preparation of ribozyme molecules for
CC inhibiting expression of the tumour polypeptides and proteins in the
CC tumour cells, in vaccines and for gene therapy.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4551 BP; 780 A; 1597 C; 1522 G; 652 T; 0 other;

Alignment Scores:
Pred. No.: 1.35e-76 Length: 4551
Score: 1019.00 Matches: 187
Percent Similarity: 89.43% Conservative: 16
Best Local Similarity: 82.38% Mismatches: 20
Query Match: 82.24% Indels: 4
DB: 24 Gaps: 1

US-09-938-391-2 (1-230) x ABV94763 (1-4551)
QY 1 ProTrpArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
DB 3874 CCTGGGGGGCAGATGACATCTCTGGCCAGCCCCCTCGCTGCCGAGCCCGAGCCCTAC 3933
QY 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
DB 3934 CCGGAGCCCGCAGCAGACCTCTACGTGACCTGGCGCGCGGCGGACCCACAGCCCA 3993
QY 41 ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuValAla 60
DB 3994 CCC-----GCCACAGCCACCGGACTTCAGCCGGTGTCCACCTCGGTGGC 4041
QY 61 LeuAsnSerProGlnProGlyGlyMetArgGlyAlaAspPheGlnCysPhe 80
DB 4042 CTCACAGCCCTGTGTGAGCGGCATCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4101
QY 81 GlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeu 100
DB 4102 CAGCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4161
QY 101 GlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeu 120
DB 4162 CAGACCTGTACAGATCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4221
QY 121 ArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlnLeu 140

```

```

Db 4222 AAGGACGAGCTGCTGTTTCCAGCTGGGAGGCTCTGTTTCTCAGGCTCTCAGGCTCGCTG 4281
QY 141 LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTrp 160
Db 4282 AAGCCCGGGGCGCGCATCTTCTCTTACGCAAGGAGCGTCTGAGGACCCACCTGG 4341
QY 161 ProArgLysSerValTrpHisGlySerAspProSerGlyArgLeuThrAspSerTyr 180
Db 4342 CCCCAGAGAGCGTGTGGCATCGCTCGGACCCACGCGGCGGCGGCGGCGGCGGCGGCGG 4401
QY 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla 200
Db 4402 TGTGAGACGTGGCGGAGCGGAGGCTCCTCGCGCCAGCGGCGGCGGCGGCGGCGGCGG 4461
QY 201 GlyArgLeuLeuGluGlnGlnAlaAlaSerCysArgHisAlaPheValValLeuCysIle 220
Db 4462 GCGAGCTCTCTGGGCGAGAGTCCGCGAGCTGCCATCACGCTTACATCGTGTCTGCATT 4521
QY 221 GluAsnSerValMetThrSer 227
Db 4522 GAGACAGCTTCATGACTGCC 4542
RESULT 4
ABN85301
ID ABN85301 standard; DNA; 4551 BP.
XX AC ABN85301;
XX 30-SEP-2002 (first entry)
XX Human collagen XVIII coding sequence.
XX Human; antirheumatic; antiarthritic; gene therapy; anti-angiogenic;
XX rheumatoid arthritis; collagen; endostatin; gene; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
FT CDS 1..4551
FT /tag= a
FT /product= "Human collagen XVIII"
XX WO200253191-A1.
XX 11-JUL-2002.
XX 03-JAN-2002; 2002WO-KR000001.
XX 05-JAN-2001; 2001KR-0000691.
XX (VIRO-) VIROMED LTD.
XX Kim J, Ho S, Park E, Kim S;
PI WPI: 2002-583596/62.
XX P-PSDB; ABN83471.
XX Novel composition for gene therapy against rheumatoid arthritis,
XX comprising a DNA encoding anti-angiogenic protein or its parts -
XX Disclosure; Page 60-70; 84pp; English.
XX The present invention relates to a composition for gene therapy,
XX comprising a DNA encoding an anti-angiogenic protein, which shows
XX therapeutic effects on rheumatoid arthritis. The composition is useful
XX for treating rheumatoid arthritis and the gene therapy is effective,
XX lasting for 14 days. The present sequence is the coding sequence for
XX human collagen XVIII. Endostatin, which consists of the C-terminal 183
XX residues of collagen XVIII, was used as an anti-angiogenic protein.
XX Sequence 4551 BP; 780 A; 1597 C; 1522 G; 652 T; 0 other;

```


KW ischaemic limb angiogenesis; diabetic neovascularisation; fracture;
 KW cytoskeletal; antiarthritic; antipsoriatic; antidiabetic; ophthalmological;
 KW synovial; gene; ss.
 OS Canis familiaris.

XX Key Location/Qualifiers
 FH 1..693
 CDS /tag= a
 FT /product= "pro-endostatin"
 FT /partial
 FT /note= "no start codon"

XX EP1191036-A2.

XX 27-MAR-2002.

XX 24-AUG-2001; 2001EP-0307224.

XX 25-AUG-2000; 2000US-227924P.

XX (PFIZ) PFIZER PROD INC.

XX Sheppard MG, Tong X;

XX WPI; 2002-354068/39.

XX P-PSDB; AAO17429.

XX An isolated nucleic acid molecule for the treatment of
 FT angiogenesis-related disorder, such as cancers or diabetic retinopathy,
 FT encodes an endostatin protein -

XX Claim 2; Fig 2; 56pp; English.

XX The present invention provides the protein and coding sequences of canine
 CC pro-endostatin and endostatin. The sequences can be used in the treatment
 CC and diagnosis of angiogenesis related disorders, including cancer,
 CC rheumatoid arthritis, psoriasis, retinopathy, macular degeneration,
 CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasia,
 CC rubesiosis, Osler-Webber Syndrome, myocardial angiogenesis, plaque
 CC neovascularisation, telangiectasia, haemophilic joints, angiodioma,
 CC wound granulation, coronary collaterals, cerebral collaterals,
 CC arteriovenous malformations, ischaemic limb angiogenesis, diabetic
 CC neovascularisation, and fractures. The present sequence is the canine
 CC pro-endostatin coding sequence.

XX SQ Sequence 829 BP; 124 A; 314 C; 278 G; 113 T; 0 other;

Alignment Scores:

Pred. No.: 3,75e-96 Length: 829
 Score: 1239.00 Matches: 230
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

US-09-938-391-2 (1-230) x AAL46062 (1-829)

QY 1 ProTTPArgAlaAspSerIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
 Db 1 CCTGGCGGGCAGATGACATCTCTGGCGGCGCGCGCTGTGGACCCCGAGCCCTAC 60
 QY 21 ProGlyValProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
 Db 61 CCGGGGCGCGCCAGCACAGCGCTCTACGTGACCTTCAGCGCGCTCGCCACTGTGGG 120
 QY 41 ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuAla 60
 Db 121 CCGGTCCACACCCACACCCACACCCAGGACTTCAGCTGTGTGACCTGTGGCC 180
 QY 61 LeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPhe 80
 Db 181 CTGAACAGCCCGAGCGGCGGCGGCGGATCCGGGAGCGGAGCTTCCAGTGTTC 240

QY 81 GlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100
 Db 241 CAGCAGCGCGCGCGCGCGCGCTGGCGGCGACCTTCGCGGCGCTTCTGCTGCGCGCTG 300
 QY 101 GlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeu 120
 Db 301 CAGGACCTCTACAGCATCGTGGCGGCGCGCGCGCGCGCGCGCTGCGTCAACCTC 360
 QY 121 ArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeu 140
 Db 361 AGGGACGAGGTGCTCTTCCCGAGCTGGAGGCGCTTATTCGGGCTCCGAGGCGCGCTG 420
 QY 141 LysProGlyValaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTyr 160
 Db 421 AAGCCCGGCGCGCGCATCTTCTTTCGACGGCAGAGATGTCTCGAGCACACCCCGCTGG 480
 QY 161 ProArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTyr 180
 Db 481 CCGCGGAGAGCGGTGTGGCAGCGCTCCGACCCCGCGGCGCGCGCTGACCGACAGCTAC 540
 QY 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuAla 200
 Db 541 TCGAGACGTGGCGGACGAGGCGCGCGCGCGCGCGCGCTGCTGCTGCTGGCG 600
 QY 201 GlyArgLeuLeuGluGlnGlnAlaAlaSerCysArgHisAlaPheValValLeuCysIle 220
 Db 601 GGCAGGCTGCTGGAGCAGGAGGCGCGCGAGCTGCCGCGCACGCTTCTGCTGCTGCTG 660
 QY 221 GluAsnSerValMetThrSerPheSerLys 230
 Db 661 GAGACAGCGTCATGACCTCTTCTCCAAG 690

RESULT 2
 ABK09977
 ID ABK09977 standard; DNA; 900 BP.
 XX
 AC ABK09977;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Synthetic plasmid pEnd-HR#1 FPD fusion DNA sequence.
 XX
 KW Mouse; Ig signal peptide; migSP; functional protein domain; FPD;
 KW primary translational product; PTP; DNA construct; regulatory DNA;
 KW DNA targeting segment; regulatory factor; single regulatory unit;
 KW monoclonal antibody; recombination-derived alteration; blood product;
 KW human; COL18A1; gene; mutant; fusion protein; ds.
 XX
 OS Chimeric - Mus sp.
 OS Chimeric - Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT exon 1..117
 FT /tag= a
 FT /number= 1
 FT /note= "Mouse Ig signal peptide (migSP). exon"
 FT CDS 61..888
 FT /tag= b
 FT /product= "pEnd-HR#1 FPD fusion protein"
 FT sig_peptide 61..117
 FT /tag= c
 FT /note= "Mouse Ig signal peptide (migSP)"
 FT mat_peptide 118..885
 FT /tag= d
 FT /note= "Human COL18A1 protein exons 38-41"
 FT exon 118..363
 FT /tag= e
 FT /number= 2
 FT /note= "Human COL18A1 exon 38"
 FT exon 364..558
 FT /tag= f

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

* Run on: August 17, 2003, 16:43:24 ; Search time 272.222 Seconds
(without alignments) 2280.750 Million cell updates/sec

Title: US-09-938-391-2.

Perfect score:

Sequence: 1 PWRADDILAGPPRLDPQY.....CRHAFVVLCIENSVMTSFSK 230

Scoring table: BLOSUM62

apop	10.0	Xgapext	0.5
apop	10.0	Ygapext	0.5
apop	6.0	Fgapext	7.0
lop	6.0	Delext	7.0

Searched: 2552756 seqs. 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```

MODEL=framed_p2n.model -DEV=xlp
C=/cgm2/1/USPTO.spool3 -Q/US09938391/runat_04082003_130655_27077/app_query.fasta_1.718
-DB=Geneseq 19JUN01 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOORCL=0
-TRANS=human40.cdci -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdci
-LOOP=45 -UNITS=bites -PCT=1 -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09938391 @CGN1_1_605 @runat_04082003_130655_27077 -NCPU=6 -ICPU=3
NO MMAP -LARGEXJECT -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database : N Genesec 19Jun03:*

1: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/cgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/cgdata/geneseq/geneseq-emb1/NA2001A.DAT
23: /SIDSI/cgdata/geneseq/geneseq-emb1/NA2001B.DAT
24: /SIDSI/cgdata/geneseq/geneseq-emb1/NA2002.DAT
25: /SIDSI/cgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	1239	100.0	829	24	AAL46062	Canine pro-endostatin	
2	1019	82.2	900	24	ABK09977	Synthetic plasmid	
3	1019	82.2	4551	24	ABV94763	Human pancreatic c	
4	1019	82.2	4551	24	ABN95301	Human collagen XVI	
5	1019	82.2	4875	24	ABQ54955	Human ovarian anti	
6	1019	82.2	5408	20	AAT77720	Human collagen 18	
7	1003	81.0	3394	18	AAT94484	Human alpha-1 coll	
8	1003	81.0	3394	20	ABX78379	Human alpha1 (XVII	
9	1003	81.0	3394	24	ABN95680	Gene #2178 used to	
10	980.5	79.1	4031	18	AAT94485	Mouse alpha-1 coll	
11	966	78.0	555	24	AAL46063	Canine endostatin	
12	959	77.4	552	21	AAS251309	Canine angiogenesi	
13	849.5	68.6	816	21	AAA64013	DNA encoding a hum	
14	846.5	68.3	641	22	ABK79104	Human endostatin e	
15	846.5	68.3	641	24	AAL54544	Specific tumour ce	
16	846.5	68.3	641	24	AAL44000	Reproductive recom	
17	846.5	68.3	641	24	ABQ76079	Anticancer gene-as	
18	843	68.0	573	21	AAC36777	Marine endostatin	
19	843	68.0	573	24	ABK47719	CDNA encoding mous	
20	842	68.0	558	24	ABA99261	Human endostatin c	
21	840.5	67.8	968	21	AAC62025	Nucleotide sequenc	
22	836	67.5	568	20	AZ087749	Mouse endostatin c	
23	836	67.5	580	20	AAZ08748	Mouse endostatin c	
24	835	67.4	546	21	AAZ37004	Human endostatin g	
25	835	67.4	549	20	AAX77719	Human endostatin D	
26	835	67.4	549	21	AAZ39884	Human angiogenesis	
27	835	67.4	549	21	AAS251291	Human angiogenesis	
28	835	67.4	549	22	AAS00867	Human gene fragmen	
29	835	67.4	549	25	ABA00774	Human endostatin c	
30	835	67.4	552	20	AAC35375	SEQ ID 50 of WO991	
31	835	67.4	552	21	AAC62033	Nucleotide sequenc	
32	835	67.4	552	21	AAA68203	Human endostatin c	
33	835	67.4	552	21	AAS250398	Human endostatin c	
34	835	67.4	552	22	AAC88289	Human endostatin c	
35	835	67.4	552	24	ABK50685	CDNA encoding huma	
36	835	67.4	555	22	AAL66529	Human vascular end	
37	835	67.4	563	20	AAZ08750	Human endostatin c	
38	835	67.4	1564	24	ABQ76740	DNA encoding human	
39	831	67.1	555	20	AAX94635	Mouse endostatin c	
40	831	67.1	555	20	AAX79949	Endostatin coding	
41	831	67.1	555	22	AAC88290	Murine endostatin	
42	831	67.1	558	22	ABZ18701	Mouse endostatin e	
43	831	67.1	558	24	ABQ76173	Rat endostatin enc	
44	831	67.1	558	24	ABQ76178	Rat endostatin enc	
45	831	67.1	565	20	AAX58740	DNA encoding anti-	

ALIGNMENTS

RESULT 1
AAL46062
ID AAL46062 standard: cDNA: 829 BP

AA
AC
AAL46062;XX
DT 19-JUL-2002 (first entry)

DE Canine pro-endostatin coding sequence.

KW Doq; pro-endostatin; endostatin; angio

KW psoriasis; pneumothorax; aortic stenopathy; macular degeneration;
KW corneal graft rejection; neovascular glaucoma; retrolental fibroplasia;
KW rubeosis; Osler-Weber Syndrome; myocardial angiogenesis;
KW plaque neovascularisation; telangiectasia; haemophilic joints;
KW angiofibroma; wound granulation; coronary collateral;
KW cerebral collateral; arteriovenous malformation;

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

Db 934 GCACTGAACACCCCTGTCGTGGAGCATGCGGTATCCGTGGAGCAGATTTCCAGTGC 993

QY 80 PheGlnAlaArgAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArg 99

Db 994 TTCACGACGAGCCGCGGGCTGTGGGACCTTCCGGCTTCTCTCTCTAGG 1053

QY 100 LeuGlnAspLeuTyrSerIleValaArgAlaAspArgThrGlyValProValValAsn 119

Db 1054 CTGCAGGATCTCTATAGCATGTCGCGCTGCTGACCGGGGCTCTGTGCCATCGTCAAC 1113

QY 120 LeuArgAspGluValLeuPheProSerTrpGluAlaIlePheSerGlySerGluGlyGln 139

Db 1114 CTGAAGCAGCAGGTGCTATCTCCAGCTGGGACTCCCTGTTTCTGGCTCCAGGTCAC 1173

QY 140 LeuLysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAla 159

Db 1174 CTGCAACCCGGGCGCCGATCTTTCTTTGAGCGCAGAGATGCTCTGAGACACCCAGCC 1233

QY 160 TrpProArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSer 179

Db 1234 TGSCCGCAGAGAGCGTATGGCAGCGCTGGACCCCGAGTGGCGGAGCTGATGGAGAGT 1293

QY 180 TyrCysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeu 199

Db 1294 TACTGTGAGCATGGCGAACTGAACTACTTGGGGCTACAGGTGAGGCTCTCCCTGCTG 1353

QY 200 AlaGlyArgLeuLeuGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCys 219

Db 1354 TCAGGAGGCTCTGGAACAGAAAGCTGGAGCTGCCAACAGCTACATCGCTGCTGTGC 1413

QY 220 IleGluAsnSerValMetThrSerPheSerLys 230

Db 1414 ATTGAGAAATGCTTCATGACCTTTCTCCAA 1446

RESULT 11

MUSCOLLAG

LOCUS

DEFINITION Mus musculus mRNA for collagen, partial cds.

ACCESSION D17546

VERSION D17546.1 GI:2160436

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE

AUTHORS Abe.N., Muragaki,Y., Yoshioka,H., Inoue,H. and Ninomiya,Y.

TITLE Identification of a novel collagen chain represented by extensive interruptions in the triple-helical region

JOURNAL Unpublished

AUTHORS Abe.N.

TITLE Direct Submission

JOURNAL Submitted (04-SEP-1993) Nobuhiro Abe, Okayama University Medical School, Molecular Biology and Biochemistry; 2-5-1, Shikata-cho, Okayama, Okayama 700, Japan (Tel:81-86-223-7151(ex.2390), Fax:81-86-222-7769)

COMMENT On Jun 5, 1997 this sequence version replaced gi:467516.

FEATURES

source

1. 4437

Location/Qualifiers

organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/dev_stage="embryo"

/note="clones NA[1,12,286,261]"

<1. 3423

/codon_start=1

/product="collagen"

/protein_id="BA044483.1"

/db_xref="GI:467517"

/translation="QRVPFARASQGLEBERGAGLFVQGQATADPKFQGMISELKVKK

TRVSPVHCLDEEDDDEDRASGDFGSPFESSKSHKEDTSLPLGLPQPPVTSPLAG

GSTTEDPRTEETEEDAAVDSIGAETLPGTSSGAWDEAIONPGRGLIKGMKGQKEP

CDS

3' UTR

polyA_signal

polyA_signal

polyA_site

BASE COUNT 962 a 1311 c 1345 g 819 t

ORIGIN

Alignment Scores:

Pred. No.: 1,126-58 Length: 4437

Score: 983.50 Matches: 185

Percent Similarity: 87.88% Conservative: 18

Best Local Similarity: 80.09% Mismatches: 23

Query Match: 79.38% Indels: 5

DB: 10 Gaps: 2

US-09-938-391-2 (1-230) x MUSCOLLAG (1-4437)

QY 1 ProTrpArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20

Db 2740 CCTCGGAGCAGATGACATCTCTGCCAACCCACCGCGCTCGCCAGCCGCGCCTTAC 2799

QY 21 ProGlyAlaPro---HisHisGlySerTyrValHisPheGlnProAlaArgProThrGly 39

Db 2800 CCGGAGTTCACATCACACAGTTCCTATGTGCCTCCGCCGCCGCCGCCACC--- 2856

QY 40 GlyProValHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuVal 59

Db 2857 -----CTCCTCCTGCTCATCTCATCAGGACTTTTCAGCCAGTGTCTCCACCTGCTG 2907

QY 60 AlaLeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCys 79

Db 2908 GCACTGAACACCCCTGTCTGGAGGATGCTGTGTATCCGTGGAGCAGATTTCAGTGC 2967

QY 80 PheGlnGlnAlaArgAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArg 99

Db 2968 TTCGAGACGCCGAGCCCTGCGGCTGTCCGACACTTCCGGCTTCTCTCTCTAGG 3027

QY 100 LeuGlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsn 119

Db 3028 CTGCAGGATCTCTATAGCATCTGCGCGCTGCTGACCGGGGCTGTGCTCCATCGTCAAC 3087

QY 120 LeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGln 139

Db 3088 CTGAAGCAGGAGTGTCTCTCCAGCTGGGACTCCCTGTTTCTGGCTCCAGGTCAC 3147

QY 140 LeuLysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAla 159

Db 3148 CTGCAACCCGGGCGCCGATCTTTCTTTGAGCGCAGAGATGCTCTGAGACACCCAGCC 3207

QY 160 TrpProArgLysSerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSer 179

Db 3208 TGGCCGACAGACAGCGTATGGCAGCGCTCGGACCCAGTGGCGGAGGCTGATGGAGAGT 3267

Percent Similarity:	88.11%	Conservative:	15
Best Local Similarity:	81.50%	Mismatches:	23
Query Match:	80.95%	Indels:	4
DB:	9	Gaps:	1
US-09-938-391-2 (1-230) x HUMCOL18AX (1-3394)			
QY	1	ProTprArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr	20
DB	1378	CCCTGGCGGCGAGATGACATCTGGCCAGCCCTCTGGGCTGCCGAGCCCGAGCCCTTAC	1437
QY	21	ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly	40
DB	1438	CCCGAGGCGCGACACACAGCTCTACGTGCTACGTGGCGCGGCGGACCCACAGCCCA	1497
QY	41	ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuValAla	60
DB	1498	CCC-----GCCACAGCCACCGGACCTTCAGCGCGTGTCCACCTGGTGTGG	1545
QY	61	LeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPhe	80
DB	1546	CTCAACAGCCCTCTTCAGCGGCGCATCGGGGCATCGCGGGCGGACCTTCAGTGTTC	1605
QY	81	GlnGlnAlaArgAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu	100
DB	1606	CAGCAGCGCGCGCGCTGGGGCTGGCGGCACCTTCGCGGCTTCTCTCGCGCTG	1665
QY	101	GlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeu	120
DB	1666	CAGGACCTGTACAGCATGTGGCGCGTCCGACCGGACCGGCTCCATCGTCAACCTC	1725
QY	121	ArgAspGluValLeuPheProSerTprGluAlaLeuPheSerGlySerGluGlyGlnLeu	140
DB	1726	AAGGACGAGCTGCTGTTTCCAGCTGGGAGGCTCTGTTCTCAGGCTCTGAGGGTCCGCTG	1785
QY	141	LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTpr	160
DB	1786	AAGCCCGGGCGACGATCTTCTCTTTCACCGCAGGACGCTCCTGAGGCACCCACCTGG	1845
QY	161	ProArgLysSerValTprHisGlySerAspProSerGlyArgArgLeuThrAspSerTyr	180
DB	1846	CCCAGAGAGCGTGTGGCATGTGCTGGACCCCAAGCGGCGAGGCTGACCGAGAGTAC	1905
QY	181	CysGluThrTprArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla	200
DB	1906	TGTGAGACGTGGCGAGCGAGGCTCCCTCGGCCACGCGGCGAGGCTCTCTGCTGGGG	1965
QY	201	GlyArgLeuLeuGluGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCysIle	220
DB	1966	GGCAGGCTCTCTGGGCGAGTGGCGGAGTGGCCATCAGCCTACATCGTCTCTGCATT	2025
QY	221	GluAsnSerValMetThrSer	227
DB	2026	GAGAACAGCTTCATGACTGCC	2046
RESULT 10			
MMU03714			
LOCUS	Mus musculus BALB/c alpha 1 type XVIII collagen (COL18A1) mRNA, partial cds.		
ACCESSION	U03714		
VERSION	U03714.1		
KEYWORDS	GI:487733		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Rehn, M. V. and Pihlajaniemi, T.		
TITLE	A previously uncharacterized collagen chain, alpha 1(XVIII), with frequent interruptions in the collagenous sequence, a distinct tissue distribution, and homology with type XV collagen		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. (1994) In press		

REFERENCE	2 (bases 1 to 1615)	Rehn, M., Hintikka, E. and Pihlajaniemi, T.	
AUTHORS		Primary structure of the alpha 1 chain of mouse type XVIII collagen, partial structure of the corresponding gene, and comparison of the alpha 1(XVIII) chain with its homologue, the alpha 1(XV) collagen chain	
JOURNAL	J. Biol. Chem. 269 (19), 13929-13935 (1994)		
MEDLINE	94245707		
PUBMED	8186673		
REFERENCE	3 (bases 1 to 1615)		
AUTHORS	Rehn, M. V.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-NOV-1993) Marko Rehn, Department of Medical Biochemistry, University of Oulu, Kajaanintie 52 A Fin-90220, Oulu, Finland		
FEATURES	Location/Qualifiers		
source	1..1615		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="BALB/c"		
	/db_xref="taxon:10090"		
	/sex="male"		
	/tissue_type="liver"		
	/clone_lib="Mouse adult liver cDNA library by M. Rehn"		
	/dev_stage="adult, 9-10 weeks"		
	<1..1449		
	/standard_name="COL18A1"		
	/citation=[2]		
	/codon_start=1		
	/product="alpha 1(XVIII) collagen"		
	/protein_id="AAA20657.1"		
	/db_xref="GI:487734"		
	/translation="DSNAFVESGRPLGQOQVQSGPKGDKGEVGPGRPGQRPID LFLHAEKMGKDGDRGAGQKGERGEPAGCGFPSSVPGPPGPGVPGIIPGKGES IRGPPGPGQPPGGIYEGRGKQPPGPPGPPSPFPGHROTIVSVGPPGPPGPPGPGMAGASAGQVRIWATYQTMLDKIREVPEGWLI FVAEREELVVRNGRKVLLLEART ALPRTGTNEVALQPLVQLHKGSPYTRREYSSTARPMWADDILANPRLDRQFPY GPVHHSSVHLHPARPTLSLAHTHQDFOPVLHLVALNPLSGGMRGIRGADQFCFQQ ARAVLSGTFRFLSRRLQDLYSIVRRADRGSPVIVNLKDEVLSPSWDLFSGSGQL QGQARIFFDGRDVLHPAWFQKSVVHSGSDPSGRRLMESYCTWRTTGTATGQASSL LSGRLEOKAASCHNSIYVLCIENSFMFSFK"		
	1450..1615		
	/citation=[2]		
	1584..1589		
	/citation=[2]		
	polya_signal		
BASE COUNT	338 a 493 c 479 g 305 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	3..88e-59	Length: 1615	
Score:	983.50	Matches: 185	
Percent Similarity:	87.88%	Conservative: 18	
Best Local Similarity:	80.09%	Mismatches: 23	
Query Match:	79.38%	Indels: 5	
DB:	10	Gaps: 2	
US-09-938-391-2 (1-230) x MMU03714 (1-1615)			
QY	1	ProTprArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr	20
DB	766	CCCTGGCGAGCATGACATCTCTGGCCAAACCCACCGCGCTGCCAGACCGCCCTTAC	825
QY	21	ProGlyAlaPro---HisHisGlySerTyrValHisPheGlnProAlaArgProThrGly	39
DB	826	CCCGAGGTCCACATCACCACAGTTCCTATGTGCACCTGCCCGCACCCGCCACCC	882
QY	40	GlyProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuVal	59
DB	883	-----CTCTCACTTGTCTCATCTCATCAGACTTTTCAGCCAGTGTCTCCACCTGGTG	933
QY	60	AlaLeuAsnSerProGlnProGlyMetArgGlyIleArgGlyAlaAspPheGlnCys	79

```
QY 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla 200
Db 1906 TGTGAGCGTGGCGGAGGAGGCTCCCTCGGCCACCGGGCCAGCCCTCTCGCTGGGG 1965
QY 201 GlyArgLeuLeuGluGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCysIle 220
Db 1966 GGCAGGCTCTCGGGCAGAGTGGCGGAGTGCATCAGCCCTACATCGTCTCTGCATT 2025
QY 221 GluAsnSerValMetThrSer 227
Db 2026 GAGACAGCTTCATGACTGCC 2046

RESULT 8
LOCUS I51045 3394 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 4 from patent US 5643783.
ACCESSION I51045
VERSION I51045.1 GI:2472748
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3394)
AUTHORS Olsen,B.R. and Oh,S.P.
TITLE Collagen and uses therefor
JOURNAL Patent: US 5643783-A 4 01-JUL-1997;
FEATURES
source Location/Qualifiers
1..3394 /organism="unknown"
BASE COUNT 668 a 1130 c 1080 g 516 t
ORIGIN

Alignment Scores:
Pred. No.: 3.83e-60 Length: 3394
Score: 1003.00
Percent Similarity: 88.11%
Best Local Similarity: 81.50%
Query Match: 80.95%
Indels: 4
Gaps: 6

US-09-938-391-2 (1-230) x I51045 (1-3394)
QY 1 ProTrpArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
Db 1378 CCTGGCGGCGACATGACATCTCTGGCCAGCCCTCTGGGCTGCCAGCCAGCCCTAC 1437
QY 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
Db 1438 CCGGAGGCGCGCACACAGCTCTACGTGCTGCGGCGCGGCGACGCCACACAGCCCA 1497
QY 41 ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuValAla 60
Db 1498 CCC-----GCCACAGCCACCGCGACTTCCAGCGGTGCTCCACCTGGTGGCG 1545
QY 61 LeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPhe 80
Db 1546 CTCAACAGCCCTCTGTCAGCGGCATCGCGGGCATCCGCGGGCCACTTCCAGTCTTC 1605
QY 81 GlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100
Db 1606 CAGCAGCGCGGCGCGTGGGGCTGGCGGCACCTTCCGCGCCCTCTGCTCTCGCGCTG 1665
QY 101 GlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeu 120
Db 1666 CAGGACCTGTACAGCATCTGTCGCCCGCTGGCGACCGCGCAGCCGTGCCATCGTCAACCTC 1725
QY 121 ArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeu 140
Db 1726 AAGGACGAGCTGTGTTTCCAGCTGGAGGCTCTGTTCTCAGGCTCTGAGGGTCCGCTG 1785
QY 141 LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTrp 160
Db 1786 AAGCCCGGGGCGACGATCTTCTCTTTGACGGCAGGACGTCCTGAGGACCCACCTGG 1845
```

```
QY 161 ProArgLysSerValTrpHisGlySerAspProSerGlyArgLeuThrAspSerTyr 180
Db 1846 CCCAGAGAGAGCGTGTGGCATGCTCGGACCCACAGGGCGGAGCTGACCGAGACTAC 1905
QY 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAla 200
Db 1906 TGTGAGCGTGGCGGAGGAGGCTCCCTCGGCCACCGGGCCAGCCCTCTCTCGCTGGGG 1965
QY 201 GlyArgLeuLeuGlnGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCysIle 220
Db 1966 GGCAGGCTCTCGGGCAGAGTGGCGGAGTGCATCAGCCCTACATCGTCTCTGCATT 2025
QY 221 GluAsnSerValMetThrSer 227
Db 2026 GAGACAGCTTCATGACTGCC 2046

RESULT 9
LOCUS HUMCOL18AX 3394 bp mRNA linear PRI 01-NOV-1994
DEFINITION Human collagen type XVIII alpha 1 (COL18A1) mRNA, partial cds.
ACCESSION L22548
VERSION L22548.1 GI:348908
KEYWORDS alpha-1 type XVIII collagen.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3394)
AUTHORS Oh,S.P., Warman,M.L., Seldin,M.F., Cheng,S.D., Knoll,J.H.,
Timmons,S. and Olsen,B.R.
TITLE Cloning of cDNA and genomic DNA encoding human type XVIII collagen
and localization of the alpha 1(XVIII) collagen gene to mouse
chromosome 10 and human chromosome 21
JOURNAL Genomics 19 (3), 494-499 (1994)
MEDLINE 94245237
PUBMED 8188291
COMMENT Original source text: Homo sapiens cDNA to mRNA.
FEATURES
source Location/Qualifiers
1..3394 /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="21q22.3"
/tissue_type="liver"
/dev_stage="adult"
1..3394 /gene="COL18A1"
<1..2055 /gene="COL18A1"
/codon_start=1
/product="collagen type XVIII alpha 1"
/protein_id="AAAS1864.1"
/db_xref="GI:562794"
/translation="GEVADGIPGPPGLPGREGIAGPOGPKDGRSGRKGDPGKDGKL
CQGLPGRGPVPVYVSEODGSVLVPGPEGRGAFGFPAGPKGNLGSKGELGS
PQPKGEKGEPSIFSPDGLGALPAOKAKAGKBEFRGPPGLYGRPGYKGEIGFGRPCR
PQNLGKGEKGEPSIFSPDGLGALPAOKAKAGKBEFRGPPGLYVDSNVFAESRRPGPG
LPNGPPGPKGKEVGPDPGQPFDFLQKEAMKEKDGDRGADGQKGRGEPGG
GFGSSLPAGAPGAPGRGPIPGPKGESINGQPPGPPQPPGPGIGYEGRQPPGP
GGPPSPFPGRHRTISVPGPPGPPGPPGPGTGMGASGVRLWATROAMLQGVHEVP
EGWLIFVAOEELVYRVONGPRKVOLEARTPLPRCTDNEVALQPPVQLHDSNPYPR
REHPHTARPNRADDILASPPGLPEPQYPPGPHSSYVHCSPARTSPPAHSRDFO
PVLHLVNLNSPLSGMGRIGADFQCFQARAVAGLAFSLSRLLQDLISIVRRAD
RAAVIVNLKDELFPISWEALFSGSEGPKPGARIFSPFDKDLVLRHPTWPKSVWHGS
DPNGRLTESYCTWRTEAPATQASLLGRLLGQSAASCHHAYIVLCIENSFMTA
SK"
BASE COUNT 668 a 1130 c 1080 g 516 t
ORIGIN

Alignment Scores:
Pred. No.: 3.83e-60 Length: 3394
Score: 1003.00
Percent Similarity: 88.11%
Best Local Similarity: 81.50%
Query Match: 80.95%
Indels: 4
Gaps: 6
```



```
gene
1..5408
/gene="COL18A1"
60..4070
/gene="COL18A1"
/translation="MAPRCPCWPRRRRLDLVLAFLVLLGLVRAASAPERISEVEVL
LQLLGGPPQVTTDDPDVGLAYVFGPDANSQVARYHFFSLFFRDFSLFHPHAT
QWGLVFAITDQAQWLVGLSGVDGHODISLITYTEPGAGQTHRAAFLRPAFVG
OWTHLALSVAAGFVALYVDCBEFQRMPLARSRLGELFEPGAGLFVAQAGADPKFQV
VIAELKVRDPQVSMHCLDGDGAFGDSGSLGDARLLREETGAALKRPLRA
PVPVTTPLAGSSSTDSERSEVERQTTVASIAGQTLPGSDSVTWDGSRVTPGERVK
EGGLKGQKEPVGVPGRAGRPFCPLGPGFLCPVSPGLGAPALQTVRPGQPP
GPPRGDTGRDGEKDPGDEKDPGDTGPGQFPGTDPVDPKDGDPGVBGRGPPGP
QQPQPGPGFGRHDLKTFIDMEGSGFGDLEALRGPPGPPGPPGPPGPPGPPGPPG
GYNSSDPVGPAGLPGVREGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
KGSKGAPGAGARGESGLAGAPGAPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
APSAQDPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
KRSRREKDPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
APGNLGSKRGSGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
PGYGEIGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
DSNVFAESRPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
RGDAGOKGSGRGGGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
PGYGEIGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
PGYGEIGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
ATRQAMLGQVHEVPEGLLFAVEBELYVRVNGRKRKQVLEARTPLRGTNDNEVALQ
PVVQLHDSNPVPRHEPHPTARPRADDILASPPRLPEPYPGAPHHSSVYHLRPA
RTPSPAHSHRDFQVPLHLVALNSPLSGMGRIGADFCQFOQARAVAGLAGTFRAFLS
SLQDLQSYIVRRADAARVIVNLKDELLFPSSWEALFSGSEGLPKPCARIFSPDGKDLV
RHPVPOKSVHMGSDPNRRLLTESYCTWRTAPSATGQASSLLGRLLGQSAASCHH
AYVILIENSPFTASK"
BASE COUNT 986 a 1820 c 1796 g 806 t
ORIGIN

Alignment Scores:
Pred. No.: 4,92e-61 Length: 5408
Score: 1019.00 Matches: 187
Percent Similarity: 89.43% Mismatches: 16
Best Local Similarity: 82.38% Indels: 4
Query Match: 82.24% Gaps: 1
DB: 9

US-09-938-391-2 (1-230) x AF018082 (1-5408)

Qy 1 ProTtpArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
Db 3393 CCCTGGCGGCGAGATGACATCTCTGGCCAGCCGCCCTCGCTGCCGAGCCCGCTAC 3452
Qy 21 ProGlyAlaProHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
Db 3453 CCGCGAGCCCGGCGACACACAGCTCTACGTGACCTTCAGCCGCGCGCGACCAAGCCCA 3512
Qy 41 ProValHisThrHisThrHisThrHisGlnAspPheGlnLeuValLeuAla 60
Db 3513 CCC-----GCCACAGCCACCGGACTTCACCGGTGCTCCACCTGGTGGC 3560
Qy 61 LeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPhe 80
Db 3561 CTCACAGCCCGCTGTGAGCGCGCATGCGGGGCATCCGCGGGCGCGGACTTCAGTGTCTC 3620
Qy 81 GlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100
Db 3621 CAGCAGCGCGCGCGCGTGGCGGTGGCGGCGACCTTCGCGCGCTTCCTGTCGCGCGCTG 3680
Qy 101 GlnAspLeuTyrSerIleValArgAlaAspAlaAspArgThrGlyValProValValAsnLeu 120
Db 3681 CAGGACCTGTACAGCATCTGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 3740
Qy 121 ArgAspGluValLeuPheProSerTtpGluAlaLeuPheSerGlySerGluGlyGlnLeu 140
Db 3741 AAGGACGAGCTGTGTTTCCAGCTGGGAGGCTCTGTTCTCAGGCTCTCAGGCTCGCGCTG 3800
```


Db 619 AAGCCCGGGGACGATCTTCTCTTGTGACGGAAGGACGCTCTGAGGACCCACCTGG 678
Qy 161 ProArglySerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTyr 180
Db 679 CCCAGAAGAGCGCTGTGGCATGCTCGGACCCCAACGGCGGAGGCTGACCGAGAGCTAC 738
Qy 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuAla 200
Db 739 TGTGACAGCTGGCGGAGGAGGCTCTCTCGGCCACGGCCAGGCTCTCTGCTGGGG 798
Qy 201 GlyArgLeuLeuGluGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCysIle 220
Db 799 GGCAGAGCTCTGGGGCAGAGTGGCGGAGTGCCTCAGCCATCAGCTCATCTGCTCTGATT 858
Qy 221 GluAsnSerValMetThrSer 227
Db 859 GAGAACAGCTTTCATGACTGCC 879
RESULT 3
LOCUS AK098216 2364 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ40897 fis, clone UTERU2002964, highly similar to Homo sapiens type XVIII collagen (COL18A1) mRNA.
ACCESSION AK098216
VERSION 1 GI:21758185
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K., Arita, M., Musashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Negahari, K., Masuho, Y., Nagai, K. and Isogai, T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2364)
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
FEATURES
source Location/Qualifiers
1. .2364
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UTERU2002964"
/tissue_type="uterus"
/clone_lib="UTERU2"
/notes="cloning vector: pME18SFL3"
BASE COUNT 420 a .768 c 776 g 400 t
ORIGIN
Alignment Scores:
Pred. No.: 2,06e-61 Length: 2364
Score: 1019.00 Matches: 187
Percent Similarity: 89.43% Conservative: 16
Best Local Similarity: 82.38% Mismatches: 20

Query Match: 82.24% Indels: 4
DB: 9 Gaps: 1
US-09-938-391-2 (1-230) x AK098216 (1-2364)
Qy 1 ProTrpArgAlaAspAspIleLeuAlaGlyProProArgLeuLeuAspProGlnProTyr 20
Db 1518 CCCTGGCGGGCAGATGACATCTCTGGCAGCCCTCCCTGGCTGCCGAGCCCGAGCTAC 1577
Qy 21 ProGlyAlaProHisHisGlySerTyrValHisPheGlnProAlaArgProThrGlyGly 40
Db 1578 CCCGAGGCCCCCACCACAGCTCTACGTGCACCTCGCGCCCTGCACGACCCCAAGCCCA 1637
Qy 41 ProValHisThrHisThrHisGlnAspPheGlnLeuValLeuHisLeuValAla 60
Db 1638 CCC-----GCCACAGCACCACCGACTTCCAGCCGGTGTCCACCTGGTTGG 1685
Qy 61 LeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPhe 80
Db 1686 CTCACAGCCCCCTGTCAGCGGCATCGGGGCATCCCGGGGGCGGACTTCAGTGTCTTC 1745
Qy 81 GlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 100
Db 1746 CAGCAGCGCGCGCGCTGGGGCTGGCGGCACCTTCCGCGCCTTCTCTCTCGCGCCTG 1805
Qy 101 GlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValProValValAsnLeu 120
Db 1806 CAGGACCTGTACAGCATCTGTGCGCCGTCGCGGAGCCGCGGAGCGCTGCTCAACCTC 1865
Qy 121 ArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeu 140
Db 1866 AAGGACGAGCTGTGTTCCAGCTGGGAGGCTCTGTTCTCAGGCTCTCAGGGTTCGCTG 1925
Qy 141 LysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHisProAlaTrp 160
Db 1926 AAGCCCGGGGACGATCTTCTTCTTTCAGGCAAGGAGCTCTCTGAGGACCCACCTGG 1985
Qy 161 ProArgGlySerValTrpHisGlySerAspProSerGlyArgArgLeuThrAspSerTyr 180
Db 1986 CCCAGAGAGCGTGTGGCATGGCTCGGACCCCAACGGCGGAGGCTGACCGAGCTAC 2045
Qy 181 CysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAlaSerSerLeuAla 200
Db 2046 TGTGAGAGCTGGCGGACGAGGCTCTCTCGGCCACCGGGCGAGGCTCTCTCGTGTGGG 2105
Qy 201 GlyArgLeuLeuGluGlnGluAlaAlaSerCysArgHisAlaPheValValLeuCysIle 220
Db 2106 GGCAGGCTCTCTGGGGCAGAGTGGCGGAGTGCCTCAGCTCAGCCCTACATCTGCTCTGATT 2165
Qy 221 GluAsnSerValMetThrSer 227
Db 2166 GAGAACAGCTTTCATGACTGCC 2186
RESULT 4
LOCUS BC033715 2637 bp mRNA linear PRI 08-JUL-2002
DEFINITION Homo sapiens, Similar to collagen, type XVIII, alpha 1, clone IMAGE:4425380, mRNA, partial cds.
ACCESSION BC033715
VERSION BC033715
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2637)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 17, 2003, 18:39:24 ; Search time 3661.11 Seconds
(without alignments)
2570.043 Million cell updates/sec

Title: US-09-938-391-2
Perfect score: 1239
Sequence: 1 PWRADILAGPRLLDQPY.....CRHAFVVLCIENSVMTSFSK 230

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DRV=xlp
-Q=/cg2_1/USPTO.spool_p/US09938391/runat_04082003_130656_27089/app_query.fasta_1.718
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938391 @CGN 1 1 5065 @runat_04082003_130656_27089 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

Result No.	Score	Query			DB ID	Description
		Match	Length	%		
1	1239	100.0	829	6	AX399629	AX399629 Sequence
2	1019	82.2	900	6	AX370851	AX370851 Sequence
3	1019	82.2	2364	9	AK098216	AK098216 Homo sapi
4	1019	82.2	2637	9	BC033715	BC033715 Homo sapi
5	1019	82.2	5408	9	AF018082	AF018082 Homo sapi
6	1019	82.2	5929	9	AF018081	AF018081 Homo sapi
7	1003	81.0	3394	6	AX409531	AX409531 Sequence
8	1003	81.0	3394	6	I51045	I51045 Sequence 4
9	1003	81.0	3394	9	HUMCOL18AX	L22548 Human colla
10	983.5	79.4	1615	10	MMU03714	U03714 Mus musculu
11	983.5	79.4	4437	10	MUSCOLLAG	D17546 Mus musculu
12	983.5	79.4	4802	10	BC043697	BC043697 Mus muscu
13	980.5	79.1	4031	6	I51044	I51044 Sequence 1
14	980.5	79.1	4031	10	MUSCOLPRO	L22545 Mouse alpha
15	966	78.0	555	6	AX399631	AX399631 Sequence
16	897.5	72.4	851	10	AF189709	AF189709 Rattus no
17	843	68.0	573	6	AR183471	AR183471 Sequence
18	843	68.0	573	6	AR183520	AR183520 Sequence
19	843	68.0	573	6	AR217279	AR217279 Sequence
20	843	68.0	573	6	AR282751	AR282751 Sequence
21	843	68.0	573	6	AX042272	AX042272 Sequence
22	843	68.0	573	6	AX490666	AX490666 Sequence
23	837	67.6	786	9	AF282883	AF282883 Homo sapi
24	835	67.4	546	6	AR193165	AR193165 Sequence
25	835	67.4	549	6	AX100086	AX100086 Sequence
26	835	67.4	552	6	AX395662	AX395662 Sequence
27	835	67.4	552	6	AX473835	AX473835 Sequence
28	835	67.4	552	6	BD081407	BD081407 Fused pro
29	835	67.4	564	9	AF416592	AF416592 Homo sapi
30	831	67.1	552	10	AF257775	AF257775 Mus muscu
31	831	67.1	555	6	BD102800	BD102800 Process f
32	831	67.1	555	6	BD128313	BD128313 Endostati
33	831	67.1	558	6	AR268609	AR268609 Sequence
34	831	67.1	565	6	AR100642	AR100642 Sequence
35	831	67.1	565	6	E34073	E34073 Carrier/DNA
36	831	67.1	624	6	AX084406	AX084406 Sequence
37	828	66.8	624	6	AX128409	AX128409 Sequence
38	822	66.3	555	9	AF184060	AF184060 Homo sapi
39	817	65.9	534	6	AR193166	AR193166 Sequence
40	817	65.9	537	6	AX100092	AX100092 Sequence
41	799	64.5	650	6	AX370853	AX370853 Sequence
42	792	63.9	5279	5	AF083440	AF083440 Gallus ga
43	754	60.9	515	10	RNO236873	Rattus no
44	736.5	59.4	5598	5	AB047066	AB047066 Xenopus l
45	730	58.9	5358	5	AY052763	AY052763 Xenopus l

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

RESULT 1

Db 82 CATACTTCCTGTATAGTTCACGTTTCATGTAATCCTCAAGAAATATAAGGAGCCAAAG 23
Qy 828 AG 829
Db 22 AG 21

Search completed: August 17, 2003, 16:22:56
Job time : 2424.32 secs

```
QY 418 CTGAAGCCCGGGCCGCGCATCTTCTTTCGACGGCAGAGATCTCTGACGACCCCGCC 477
Db 459 CTGAAGCCCGGGCCGCGCATCTTCTTTCGACGGCAGAGATCTCTGACGACCCCGCC 400
QY 478 TGCGCCCGGAGAGCGGTGGACGGCTCCGACCCCGAGCGGGCCGCTGACGACGAGC 537
Db 399 TGCGCCCGGAGAGCGGTGGACGGCTCCGACCCCGAGCGGGCCGCTGACGACGAGC 340
QY 538 TACTGCGAGACGTGGCGGACGGAGCGCCCGCGGCCACCGGGCAGGCGTCTGCTGCTG 597
Db 339 TACTGCGAGACGTGGCGGACGGAGCGCTCCGCGGACGGCCGAGGCTCTGCTGCTG 280
QY 598 GCGGCGAGGCTGTGAGCAGGAGCGCGGAGCTGCGGCGCACCGCTTCGTGCTGCTG 657
Db 279 GCGGCGAGGCTGTGCGGACGAGTGGCGGAGTGCATCAGCCCTACATCGTCTGCTG 220
QY 658 ATCGAGAACGCTCATGACCTCTCTTCAAGTAGGGCGCGCGGCCACGAGAGGCG 717
Db 219 ATTGAGAACAGCTTCATGACTGCTCCCAAGTAGCCACCGCTGGATGCGGATGCGCGAG 160
QY 718 GGGGAGGGCGCGCGCAGGAGCATCGCGCGCCCGCGGG-----756
Db 159 AGGACCGCGGCTCGGAGGAGCCCGACCGTGGGAGGAGCGCGCGCCCGCTGG 100
QY 757 -----GGGCGTGGCGGAGCGCTTGCTGTCACCG-TCACGTTTAAATGTAATCTCAAGAA 810
Db 99 CCCCAGGACCTGCTGCCATCTTCTGTATAGTTCAAGTTTCATGTAATCTCAAGAA 40
QY 811 ATAAAGGAGCAAGAG 829
Db 39 ATAAAGGAGCAAGAG 21
```

```
RESULT 15
BM683067/c
LOCUS
DEFINITION
  BM683067 683 bp mRNA linear EST 27-FEB-2002
  UI-E-EO1-ajd-g-02-0-UI.s1 UI-E-EO1 Homo sapiens cDNA clone
  UI-E-EO1-ajd-g-02-0-UI 3', mRNA sequence.
ACCESSION
  BM683067
VERSION
  BM683067.1 GI:18992963
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 683)
REFERENCE
  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
JOURNAL
  9704477
MEDLINE
  8889548
PUBMED
  Contact: Soares, MB
  Coordinated Laboratory for Computational Genomics
  University of Iowa
  375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
  Tel: 319 335 8250
  Fax: 319 335 9565
  Email: bento-soares@uiowa.edu
  Tissue Procurement: Dr. Gregg Hageman
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Researchers may obtain clones from Research
  Genetics (www.resgen.com).
  The following repetitive elements were found in this cDNA
  sequence: 1-44, >POLY A#simple_repeat (matched complement)
  Seq primer: M13 Forward
  POLYA=Yes.
FEATURES
  source
    1..683
    /organism="Homo sapiens"
    /mol_type="mRNA"
```

```
/db_xref="taxon:9606"
/clone="UI-E-EO1-ajd-g-02-0-UI"
/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EO1"
/notes="Organ: eye; Vector: p77T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-EO1 is a normalized cDNA library containing the
following tissue(s): fetal eye. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into p77T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dt)18 tail. The sequence tag for this library is
CGCGTATACC. This library was created for the program, Gene
Discovery in the Visual System, supported by National Eye
Institute (NEI).
TAG_LIB=UI-E-EO1
TAG_TISSUE=human fetal eye
TAG_SEQ=CGCGTATACC"
BASE COUNT 112 a 217 c 221 g 132 t 1 others
ORIGIN
Query Match 46.7%; Score 387; DB 12; Length 683;
Best Local Similarity 76.9%; Pred. No. 1.7e-64;
Matches 509; Conservative 0; Mismatches 126; Indels 27; Gaps 2;
QY 195 GCGGCGGCGCATCGGAGGCGATCCGGGAGCGACTTCCAGTGTCTCCAGAGCGCGCGC 254
Db 682 GTCAGCGCGCATCGGCGGCGATCCGGGCGCGACTTCCAGTGTCTCCAGAGCGCGCGC 623
QY 255 CCGCGGCGTGGCGGCGACCTTCGCGGCGCTTCCTGTCTGTCGCGGCTGCAGGACCTCTACAG 314
Db 622 CGTGGGCGTGGCGGCGACCTTCGCGGCGCTTCCTGTCTGTCGCGGCTGCAGGACCTGTACAG 563
QY 315 CATGTGCGCGCGCGCGGCGACCGGCGTGGCGGCGCTTCCTGTCTGTCGCGGCTGCAGGACGAGTGTCT 374
Db 562 CATGTGCGCGCGTGGCGGCGACCGGCGCGCTTCCTGTCTGTCGCGGCTGCAGGACGAGTGTCT 503
QY 375 CTTCCCGCGTGGGAGCGCTTATTCGCGGCTCCGAGGCGCGAGCTCAAGCGCGGCGCGC 434
Db 502 GTTTCGCGTGGGAGCGCTGTGTCTGAGGCTCTGAGGCTCGCGTCAAGCGCGGCGCGC 443
QY 435 CATCTTCTCTTCGAGCGGCGAGATGTCTGCGAGCACCCCGCGCTGCGCGCGCGGAGAGCGT 494
Db 442 CATCTTCTCTTCGAGCGGCGAGCGTCTCTGAGGCGACCCCACTGCGCGCGGAGAGCGT 383
QY 495 GTGGCAGCGCTCCGACCCCGCGGCGCGCTGACCGAGAGTACTGCGAGACGTTGGCG 554
Db 382 GTGGCATGCTCGGACCCCGCGGCGCGAGGCTGACCGAGAGTACTGTGAGACGTTGGCG 323
QY 555 GACGAGCGCGCGGCGCGCGCGCGCGCGCTGCTGCTGCGCGGCGGCGGCGTCTGGA 614
Db 322 GACGAGGCTCTCTTCGCGCGCGCGCGCGCGCTCTCTGCTGCGGCGGCGGCGTCTTGGG 283
QY 615 GCAGAGGCGCGGAGCTGCGCGCGCGCGCTTCGTGCTGCTGTCGATCGAAGACGCTCAT 674
Db 262 GCAGAGTGGCGGAGTGCATCAGCGCTACATCGTGTCTGCTGCTGATGAGACGCTTCT 203
QY 675 GACCTCTTCTCAAGTAGGGCGCGCGCGCGCGCGCGAGCGAGCGGCGGCGCGCGCGC 734
Db 202 GACTGCTCCAAAGTAGCCACCGCTGGATGTCAGTGGCGCGAGAGACCGCGGCTCGGA 143
QY 735 AGGAGCATCGCGCGCGCGCGCGG-----GGGCGTGGCGCG 768
Db 142 GGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 83
QY 769 GACGCTTGTCTGCGACCG-TCACGTTTAAATGTAATCTCAAGAAATAAAGGAGCAAG 827
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6650260"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 109"
/notes="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."
BASE COUNT 163 a 260 g 190 t
ORIGIN

Query Match 48.0%; Score 398.2; DB 13; Length 944;
Best Local Similarity 78.5%; Pred. No. 1.2e-66;
Matches 489; Conservative 0; Mismatches 133; Indels 1; Gaps 1;

QY 139 CACACCACAGGACTCCAGCTGGT-GCTGCACCTGTGGCCCTGAACAGCCCGCAGCC 197
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
712 CAGTTCGACCGAGGCTCCAGCGGGTGGTCCACATGGTTGGCTCAGCAGCCCCCTGTC 653
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
198 GGGCGGCATCGCAGGCATCCGGGAGCGGACTTCCAGTGCTTCCAGCAGCGCGCGCCGC 257
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
652 AGCGGCATGCGGGGATCCGCGGGCCGACTTCCAGTGCTTCCAGCAGCGCGCGCGGT 593
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
258 GGGGTGGCGGCACCTTCCGGGCTTCTGTGCTCGCGGCTCCAGGACCTTACAGCAT 317
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
592 GGGGTGGCGGCACCTTCCGGGCTTCTGTGCTCGCGGCTCCAGGACCTTACAGCAT 533
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
318 CFTGCGCGCGCGCAGCGGACCGGGTGGCGGCTCGTCACTCAGGACGAGTGTCTTT 377
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
532 CGTGGCGCGCGCAGCGGACCGCGGCTCGTCACTCAGGACGAGTGTCTTT 473
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
378 CCCAGCTGGGAGGCTTATTTCTGGGCTCCGAGGGCCAGCTGAAGCCCGGGCCCGCAT 437
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
472 TCCAGCTGGGAGGCTCTGTTCTAGGCTCTGAGGCTCGCTGAAGCCCGGGCAGCAT 413
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
438 CTCTCTTTTCAGCGCAGAGATGCTTCAGACACCCCGCTGCGCCCGGGAAGCGGTG 497
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
412 CTCTCTTTTACGGCAAGACGCTCTGAGGACACCCACCTGCGCCCGGGAAGCGGTG 353
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
498 GCACGGCTCCGACCCAGCGGCGCGCTGACCGACAGTACTGCGAGCTGTGGCGGAC 557
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
352 GCATGGCTCGGACCCCAACCGGCGCAGGCTGACCGAGAGTACTGTGAGACGTGGCGGAC 293
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
558 GGAGGGCCCGCGGACACCGGCGAGGCGTCTGCTGCTGGCGGCGAGGCTGTGGAGCA 617
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
292 GGAGGCTCTCTCGGCCACGGCCAGGCGCTCTGCTGCTGGGGGCGAGGCTCTGGGGCA 233
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
618 GGAGGCGCGAGCTGCGGCCACCGCTTCTGTTGTTGCTCTGATCGAGAACAGCGTATGAC 677
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
232 GAGTGGCGCGAGCTGCCATCAGCGCTACATCGTCTCTGATTTGAGAACAGCTTATGAC 173
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
678 CTCTTTTCAAGTAGGCGCGCGGCGCCAGGACAGCGGGGAGGGGCGCGCCGAGG 737
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
172 TGGCTTCAAGTAGCACACCGCTTGATGCATGATGGCCGAGAGGACCGGGCGGCTCGGAGGA 113
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
738 AGATTCGCGCCCGCGGGGGG 760
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
112 AGCCCCCAGCGTGGGCAAGGAGC 90
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
BU615520/c
LOCUS BU615520 703 bp mRNA linear EST 23-SEP-2002
DEFINITION UI-H-FG0-bcp-p-10-0-UI.s1 NCI_CGAP_EN1.2 Homo sapiens cDNA clone
UI-H-FG0-bcp-p-10-0-UI 3', mRNA sequence.
ACCESSION BU615520
VERSION BU615520.1 GI:23281735
```

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 703)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-44, >POLY A#simple_repeat (matched complement)
Seq primer: M3 FORWARD
POLYA=Yes.

FEATURES
source
1..703
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FG0-bcp-p-10-0-UI"
/tissue_type="Enchondroma cell line"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP EN1.2"
/note="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI_CGAP_EN1.2 is a cDNA library containing the following
tissue(s): Enchondroma cell line (2 cell lines). The
library was constructed according to Bionaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is CCGTCACTC. The cell lines was provided by Dr
James Martin from University of Iowa.
TAG_L1B=UI-H-FG0
TAG_TISSUE=Enchondroma cell line (Mix of EN1 and EN2)
TAG_SEQ=CGTCACTC"
BASE COUNT 116 a 224 c 228 g 134 t 1 others
ORIGIN

Query Match 47.6%; Score 394.4; DB 13; Length 703;
Best Local Similarity 76.6%; Pred. No. 6.6e-66;
Matches 520; Conservative 0; Mismatches 132; Indels 27; Gaps. 2;
QY 178 GCCTCAACAGCCCGCAGCGCGGCGCATCGCAGGATCGGGAGCGGACTTCCAGTGC 237
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
699 GCGCTCAACAGCCCCCTGTGACGGCGCATCGCGGGCATCGCGGGCGGCGACTTCCAGTGC 640
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
238 TTCACGACGCGCGCGCGCGGCTGGCGCGGACCTTCGCGGCTTCTGTGCTGCGCGG 297
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
639 TTCACGACGCGCGCGCGGCTGGCGGCTGGCGGACCTTCGCGGCTTCTGTGCTGCGCG 580
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
298 CTGACGAGCTTACAGCATCGTGC CGCGCGCGACCGACCGGAGTGCCTGTCGTCAC 357
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
579 CTGACGAGCTTACAGCATCGTGC CGCGCGCGACCGACCGGAGTGCCTGTCGTCAC 520
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
358 CTGACGAGCTTACAGCATCGTGC CGCGCGCGACCGACCGGAGTGCCTGTCGTCAC 417
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
519 CTCAGGACGAGCTGCTGTTTCCAGCTGGGAGGCTCTGTTCTCAGGCTCTGAGGCTCG 460

Db 304 CTCAACAGCCCCCTGTACAGCGGCATCCGGGGCATCCGGGGGCCGACTTCCAGTGCCTTC 363
QY 241 CAGCAGCGCGCGCCGGGGCTGGCGGCACCTTCGGGGCCCTTCCTGTCTGTCGGGGCTG 300
Db 364 CAGCAGCGCGCGCCGGGGCTGGCGGCACCTTCGGGGCCCTTCCTGTCTGTCGGGGCTG 423
QY 301 CAGGACCTCTACAGCATCGTGGCGCGCGCCGACCGCACCGGGGTGCCCGTCTGTCACCTTC 360
Db 424 CAGGACCTCTACAGCATCGTGGCGCGCGCCGACCGCACCGGGGTGCCCGTCTGTCACCTTC 483
QY 361 AGGACAGGTGCTCTTCCCGAGCTGGAGGCCCTTATTCCTGGGGTCCGAGGGCCAGCTG 420
Db 484 AAGGACAGCTGCTGTCTTCCAGCTGGAGGCTCTGTCTCAGGCTCTGAGGGTCCGCTG 543
QY 421 AAGCCGGGGCCGCACTCTCTCTTCGACGGCAGAGATGCTCTGAGCACCCCGCTGG 480
Db 544 AAGCCGGGGCCGCACTCTCTCTTCGACGGCAGAGATGCTCTGAGGGCACCCCACTGG 603
QY 481 CCCCAGGAGAGCGTGTGGCAGCGCTCCGACCCCGAGGGGGCGCGCTGACCCAGCAGCTAC 540
Db 604 CCCCAGGAGAGCGTGTGGCAGCGCTCCGACCCCGAGGGGGCGCGCTGACCCAGCAGCTAC 663
QY 541 TCGAGACGTGGCGGACGAGGCCCGCGGGCCCGCACCGGGCA-GGCTGTCTGCTGTGCTG 599
Db 664 TGTGACACGTGGCGGACGAGGCTCCCTCGGCACCGGGCAGGGCCCTCTGCTGTGCTGG 723
QY 600 GGG 602
Db 724 GGG 726

RESULT 9
BE908201
LOCUS 601500458F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3902175 5',
DEFINITION mRNA sequence.
ACCESSION BE908201
VERSION BE908201.1 GI:10402537
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 707)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM9704 row: d column: 16
High quality sequence stop: 688.
Location/Qualifiers
1. .707
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3902175"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 70"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 Kb. Library constructed by Life
Technologies."

FEATURES
source

BASE COUNT 113 a 250 c 225 g 119 t
ORIGIN

Query Match 51.0%; Score 422.8; DB 10; Length 707;
Best Local Similarity 78.2%; Pred No. 2.4e-71;
Matches 508; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
QY 129 CACCCACACCCACACCCACAGGACTTCAGCTGGTGTGACCTTGGTGGCCCTGAACAG 188
Db 1 CCCACCCGCCACACGCCACCGGACTTCCAGCCGGTGTCTCCACCTGGTGGCTCAACAG 60
QY 189 CCGGACGCGGGCGGCATGCGAGGATCCGGGGAGGGAGCTTCCAGTGTCTCCAGCAGGC 248
Db 61 CCCCCTGTGACGCGCATGCGGGGCATCCGGGGGCCGACTTCCAGTGTCTCCAGCAGGC 120
QY 249 GGGCGCGCGGGCTGGCGGCACCTTCGCGGCCCTTCTGTCTGTCGCGCTGCAGACCT 308
Db 121 GGGGCGCGTGGGCTGGGGGACCTTTCGCGGCCCTTCTGTCTGCGGCTGCAGACCT 180
QY 309 CTACAGCATCGTGGCGCGCGCCGACCGCACCGGGGTGCCCGTCTCAACCTCAGGACGA 368
Db 181 GTACAGCATCGTGGCGCGCGTGGCGACCGCGCAGCCCGTGCCTCACTCAAGGACGA 240
QY 369 GGTGCTCTTCCCGAGCTGGAGGCTTATTCGCGGCTCCGAGGGCCAGCTGAAGCCCG 428
Db 241 GCTGCTGTTCACAGCTGGAGGCTCTGTCTCAGGCTCTGAGGGTCCGCTGAAGCCCG 300
QY 429 GSCCGCATCTTCTCTTTCGACGCGAGATGCTCTGACGACACCCCGCTGCCCCGGAA 488
Db 301 GGCAGCATCTTCTCTTTCGACGCGAGATGCTCTGAGGACCCCGCTGCCCCAGAA 360
QY 489 GAGCGTGTGGCAGCGCTCCGACCCCGAGGGGGCGCGCTGACCGACAGCTACTGCGAGAC 548
Db 361 GAGCGTGTGGCATGCTCGGACCCCAACGGGCGCAGGCTGACCGAGAGCTACTGTGAGAC 420
QY 549 GTGGCGGACGAGGGCCCCCGGGCCACCGGGCAGGGTGTCTGCTGCGGGGAGGCT 608
Db 421 GTGGCGGACGAGGGCTCCCTCGGCCACCGGCCAGGCTCTCTGCTGCTGGGGGAGGCT 480
QY 609 GCTGGACGAGGAGCGCGGAGCTGCGCCACCGCTTCTGTTGGTGTCTGCTGCTGAGAACAG 668
Db 481 CTTGGGCGAGAGTCCGCGAGCTGCATCACGCTTACATCGTGTCTGTGCTGTGAGAACAG 540
QY 669 CTTATGACCTCTCTTCTTCCAGTAGGGCGCGCGCCACCGACAGGGGGGAGGGGCG 728
Db 541 CTTATGACTGCTCTCAAGTAGCCACCGCTGTGATGTCAGATCGCGGAGAGACCGGG 600
QY 729 GCCCGCAGAGATCCGCGGCCCGGGGGGCTGCGCGGAGCGCTTGGC 778
Db 601 CTCGAGGAAGCCCAACGTTGGGCGAGGAGCAGCGCGCCCGCTTGGC 650

RESULT 10
BE906253
LOCUS 601502237F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3904208 5',
DEFINITION mRNA sequence.
ACCESSION BE906253
VERSION BE906253.1 GI:10399595
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 757)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM9704 row: d column: 16
High quality sequence stop: 688.
Location/Qualifiers
1. .707
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3902175"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 70"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 Kb. Library constructed by Life
Technologies."

AUTHORS Ota.T., Wakamatsu.A., Ozawa.M., Ishii.S., Saito.K., Yamamoto.J., Nakamura.Y., Nishikawa.T., Nagai.T., Suzuki.Y., Sugano.S. and Isogai.T.

TITLE HRI human cDNA project (Ota.T., Wakamatsu.A., Ozawa.M., Ishii.S., Saito.K., Yamamoto.J., Nakamura.Y., Nishikawa.T., Nagai.T., Suzuki.Y., Sugano.S., Isogai.T.)

JOURNAL Unpublished

COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp

HRI human cDNA project: 5' - & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES
source
1. .715
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RM4001897"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_lib="NT2RM4"
/notes="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal precursor cells"

BASE COUNT 111 a 278 c 220 g 103 t 3 others

ORIGIN
Query Match 51.4%; Score 425.8; DB 9; Length 715;
Best Local Similarity 82.3%; Pred. No. 6.4e-72;
Matches 506; Conservative 0; Mismatches 97; Indels 12; Gaps 1;

QY 1 CCTGGCGGGCAGATGACATCTCTGGCGGCCCCCGCGCTGTGGACCCCGACCTAC 60
DB 99 CCTGGCGGGCAGATGACATCTCTGGCGGCCCCCGCTGTGGACCCCGACCTAC 158
QY 61 CCGGGGGCCCGACACAGGCTCTAGTGTGACCTTCAGCGGCTGCCCCACTGTGGG 120
DB 159 CCGGAGCGCCCGACACAGCTCTAGTGTGACCTTCAGCGGCGGACCGACCAAG--- 214
QY 121 CCGGTCCACACCCACACCCACACCCAGGACTTCAGCTGTGTGTGACCTGTGGCC 180
DB 215 -----CCACCGGCGGACAGCCAGGACTTCAGCGGCTGTCCACCTGTGTGCG 266
QY 181 CTGAACAGCCCGAGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
DB 267 CTCAACAGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 326
QY 241 CAGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 327 CAGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 386
QY 301 CAGGACCTCTACAGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
DB 387 CAGGACCTGTACAGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 446
QY 361 AGGGACAGGTGTCTTCCCGAGCTGGGAGGCGCTTATTCTGGGCGTCCGAGGGCGGAGCTG 420
DB 447 AAGGACAGGTGTCTTCCCGAGCTGGGAGGCGCTTATTCTGGGCGTCCGAGGGCGGAGCTG 506
QY 421 AAGCCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
DB 507 AAGCCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 566
QY 481 CCGCGAAGAGCGGTGTGGCAGGCTCCGACCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
DB 567 CCCAGAGAGCGGTGTGGCAGGCTCCGACCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 626
QY 541 TGCAGACGTGGGCGGAGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600

Db 627 TGTGACAGCTGGCGGACGAGGCTCCCTCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGG 686

QY 601 GCAGGCTCTGTGGAG 615

Db 687 GGCAAGCTTCTTGGG 701

RESULT 8
BUS556872 947 bp mRNA linear EST 16-SEP-2002
LOCUS AGENCOURT_10186760 NIH_MGC_109 Homo sapiens cDNA clone
DEFINITION IMAGE:6584644 5', mRNA sequence.
ACCESSION BUS556872
VERSION BUS556872.1 GI:22907168
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 947)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2795 row: f column: 04
High quality sequence stop: 616.

FEATURES
source
1. .947
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6584644"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 109"
/notes="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT 163 a 353 c 311 g 120 t

ORIGIN
Query Match 51.4%; Score 425.8; DB 13; Length 947;
Best Local Similarity 84.2%; Pred. No. 6.3e-72;
Matches 508; Conservative 0; Mismatches 82; Indels 13; Gaps 2;

QY 1 CCTGGCGGGCAGATGACATCTCTGGCGGCCCCCGCGCTGTGGACCCCGACCTAC 60
DB 136 CCTGGCGGGCAGATGACATCTCTGGCGGCCCCCGCTGTGGACCCCGACCTAC 195
QY 61 CCGGGGGCCCGACACAGCTCTACGTGCTTCAGCGGCTCGCCCCACTGTGTGGG 120
DB 196 CCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 251
QY 121 CCGGTCCACACCCACACCCACACCCAGGACTTCAGCTGTGTGTGACCTGTGGCC 180
DB 252 -----CCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 303
QY 181 CTGAACAGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240

```
QY 181 CTGAACAGCCCGCAGCGCGGCGGCATCGAGGATCCGGGAGCGGAGCTTCCAGTCTTC 240
Db 256 CTCAACAGCCCGCCTGTGTCAGCGGCATCGCGGCATCGCGGCGCCAGCTTCCAGTCTTC 315
QY 241 CAGCAGCGCGCGCGCGGCGCTGGCGGCACCTTCGGGCGCTTCTGCTGCTCGCGGCTG 300
Db 316 CAGCAGCGCGCGCGCGGCGCTGGCGGCACCTTCGGGCGCTTCTGCTGCTCGCGGCTG 375
QY 301 CAGGAGCTCTACAGCATCGTGCGCGCGCGCGCACCGGCGTCCGCTGCTCAACCTC 360
Db 376 CAGGAGCTGTACAGCATCGTGCGCGCGCGCGCACCGGCGTCCGCTGCTCAACCTC 435
QY 361 AGGAGCAGGCTGCTTCTCCAGCTGGGAGGCGCTTATTCGCGGCTCCGAGGCGCAGCTG 420
Db 436 AAGGAGCAGCTGCTGTTTCCAGCTGGGAGGCTTCTGCTCAGGCTCTGAGGCTCGGCTG 495
QY 421 AAGCCCGGCGCGCGCATCTTCTTTCGAGGCGAGAGTCTGCTGAGCACCGCGCTGG 480
Db 496 AAGCCCGGCGCGCGCATCTTCTTTCGAGGCGAGAGTCTGCTGAGGCGAGGCTGCTGAGGCGAGGCTG 555
QY 481 CCCCGAAGAGCGTGTGGCAGCGCTCCGACCCCGCGCGCGCGCGCGCGCGCGCGCGCG 538
Db 556 CCCCGAAGAGCGTGTGGCAGCGCTCCGACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 615
QY 539 ACTCGAGAGCTGGCGGAGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 598
Db 616 ACTGTGAGAGCTGGCGGAGCGAGGCTCCCTCGGACAGCGCGCGCGCGCGCGCGCGCG 674
QY 599 CGGCGAGGCTGTGGAGAGGAGGCGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 658
Db 675 GGGGCGAGGCTCTGGGCGAGGTCGCGAGAGCTTGCCTCAGCGCTACATCGTCTGCA 734
QY 659 TCGAGAACAG 668
Db 735 TTGGAACAG 744
```

```
RESULT 6
BF385854
LOCUS 602046021F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:4195660 5',
DEFINITION mRNA sequence.
ACCESSION BF385854
VERSION BF385854.1 GI:11367159
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 843)
JOURNAL NIH-MGC http://mgs.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LL9531 row: i column: 05
High quality sequence stop: 761.
FEATURES
source
1. .843
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4195660"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Li9"
```

```
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 167 a 278 c 228 g 170 t
ORIGIN
Query Match 51.6%; Score 428; DB 10; Length 843;
Best Local Similarity 77.9%; Pred. No. 2.4e-72;
Matches 547; Conservative 0; Mismatches 140; Indels 15; Gaps 2;
QY 1 CCTGGCGGCGCAGATGACATCTCTGCGCGCGCGCGCGCGCTTGGACCCCGAGCCCTAC 60
Db 42 CCTGGCGGCGCAGATGACATCTCTGCGCGCGCGCGCGCGCTTGGACCCCGAGCCCTAC 101
QY 61 CCGGGGG---CCCGCCAGCAGCGCTCTACGTGCACTTCAGCGCGCTCGCGCCACTGGT 117
Db 102 CCGGAGTTCCACATCACCACAGTTCTATGTCACCTGCGCGCGCGCGCGCGCGCGCGCG 157
QY 118 GGGCGCGTCCACACCCACACCCACAGGACTTCCAGTGTGCTGCACCTGGTG 177
Db 158 -----CCTCTCATTTGCTCATCTATCAGGACTTTCAGCAGTGTCTCACCTGGTG 209
QY 178 GCGCTGAACAGCCCGCAGCGCGCGCGCATGCGAGGCGATCCGGGAGCGGACTTCCAGTGC 237
Db 210 GCACTGAACACACCCCTGTCTGAGGCGATGCGTGGTATCCGTGGAGCAGATTTCCAGTGC 269
QY 238 TTCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 297
Db 270 TTCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 329
QY 298 CTGAGGAGCTCTACAGATCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 357
Db 330 CTGAGGAGCTCTATAGCATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 389
QY 358 CTGAGGAGCGAGTGTCTTCCCGAGCTGGGAGGCGCTTATTCGCGGCTCCGAGGCGCG 417
Db 390 CTGAAGGAGCGAGTGTCTATCTCCAGCTGGGACTCCCTGTTTCTGGCTCCCGAGGTC 449
QY 418 CTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 477
Db 450 CTGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 509
QY 478 TGGCGCGCGAGAGCGTGTGGCAGCGCTCCGACCCCGCGCGCGCGCGCGCGCGCGCG 537
Db 510 TGGCGCGAGAGAGCGTATGGACCGCTCGGACCCCGAGTGGCGGCGGCGTGTGAGAGT 569
QY 538 TACTCGGAGCGTGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
Db 570 TACTGTGAGACATGCGGAACCTGAAACTACTGGGCGCTACAGGTCAAGGCTCCCTCG 629
QY 598 GGGCGCGGCTGTGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657
Db 630 TCAGGCGGCTCTCGGAAAGAGTGGAGCTGCGAGCTGCCACAGAGTACATCGTCTGTGC 689
QY 658 ATCGAGAGCAGGCTCATGACCTCTTCTCCAGTAGGCGCGCG 699
Db 690 ATTGAAGATAGCTTCATGACCTCTTCTCCAAATAGGCGCTC 731
RESULT 7
LOCUS AU125614
DEFINITION AU125614 NT2RM4 Homo sapiens cDNA clone NT2RM4001897 5', mRNA
sequence.
ACCESSION AU125614
VERSION AU125614.1 GI:10950330
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 715)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```


W. M. Keck Center for Comparative and Functional Genomics
University of Illinois at Urbana-Champaign
340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL
61801, USA
Tel: 217 333 5998
Fax: 217 244 5817
Email: h-lewin@uiuc.edu
Funding for cattle EST sequencing was provided by the USDA National
Research Initiative, Animal Genome Resource Grant AG 99-3205-8534
to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED
from Washington University Genome Center. Vector Trimmer g.
Cross_match from Washington University Genome Center PHRAP suite.
Sequences submitted are vector free and at least 200 bp in length.
PCR Primers
FORWARD: TAATACGACTCCTACTATAGG
BACKWARD: ATTAACCTCTCACTAAG
Insert Length: 279 Std Error: 0.00
Plate: BP230015B10 row: C column: 11
Seq primer: AGCGATAACAATTCACACAGGA
High quality sequence stop: 279.

FEATURES

source
1. 279
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="BP230015B10C11"
/sex="female"
/lab_host="DH10B"
/clone_lib="Soares normalized bovine placenta"
/note="Organ: placenta; Vector: p773Pac; Site 1: EcoRI;
Site 2: NotI; The cDNA library was contributed by the
Soares laboratory and it was constructed and normalized
as described by Bonaldo, M.F., Lennon, G. and Soares,
M.B. (1996), Genome Research 6(9): 791-806."

BASE COUNT 35 a 106 c 85 g 52 t

ORIGIN

Alignment Scores:
Pred. No.: 6,41e-30 Length: 279
Score: 43.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.70% Indels: 0
DB: 9 Gaps: 0

US-09-938-391-2 (1-230) x AW464343 (1-279)

QY 116 ProValValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGly 135
|||
DB 138 CCGTCGCTCAACCTCAGGACGAGGTGCTGTTCTTAGTGGAGGCTTGTCTCAGGC 197
|||
QY 136 SerGluGlyGlnLeuLysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeu 155
|||
DB 198 TCCGAGGGCCAGCTGAAGCCGGCCGCATCTTCTCTTCACGCGCAGAGATGCTCT 257
|||
QY 156 GlnHisPro 158
|||
DB 258 CAGCATCCC 266
|||

RESULT 4

BF074459
LOCUS 221893 MARC 2BOV Bos taurus cDNA 5', mRNA linear EST 25-APR-2001
DEFINITION BF074459
ACCESSION BF074459.1 GI:10867970
VERSION
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

EST. Bos taurus

ORGANISM

Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 551)
Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,

REFERENCE

AUTHORS

TITLE

JOURNAL
MEDLINE
PUBMED
COMMENT

Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett
, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,
Perteau, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and
Keefe, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)

Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACACGCTATGACCAT
BACKWARD: GTTTCAGTCAGCAGC
Plate: 81 row: F column: 9
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers
1. 551
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 2BOV"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, thymus,
semitendinosus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."

BASE COUNT 80 a 200 c 179 g 92 t

ORIGIN

Alignment Scores:
Pred. No.: 1.3e-29 Length: 551
Score: 43.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.70% Indels: 0
DB: 10 Gaps: 0

US-09-938-391-2 (1-230) x BF074459 (1-551)

QY 70 ArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaGlyLeuAla 89
|||
DB 46 CGCGGCATCCGCGGCGCGGCTTTCAGTCTCCAGCAGCGCGCGCGGGCTGGCC 105
|||
QY 90 GlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyfSerIleValArgArg 109
|||
DB 106 GGCACCTTCGCGGCTTCTGCTCTCGCGGTTCAGACCTGTACAGCATCGTGCCTCCG 165
|||
QY 110 AlaAspArg 112
|||
DB 166 GCCGACCGT 174
|||

RESULT 5

CB424313/c
LOCUS CB424313
DEFINITION 598557 MARC 6BOV Bos taurus cDNA 3', mRNA linear EST 25-MAR-2003
ACCESSION CB424313
VERSION CB424313.1 GI:29194232
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 623)
REFERENCE

AUTHORS Smith,T.P.L., Roberts,A.J., Echtenkamp,S.E., Chitko-McKown,C.G.,
 Wray,J.E. and Keele,J.W.
TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL Unpublished
COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called with phred v0.020425.c and
 trimmed with the aid of the trim_alt option. Vector identified with
 cross_match v0.990329.
 Plate: FQY8006 row: I column: 12
 Seq primer: TAGAAGGCACAGTCGAGG.
FEATURES Location/Qualifiers
 source
 1..623
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 6BOV"
 /note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
 Library made with RNA pooled from multiple tissues
 including liver, lung, hypothalamus, pituitary, and
 placenta/endometrium."
BASE COUNT 105 a 199 c 193 g 126 t
ORIGIN
 Alignment Scores:
 Pred. No.: 1.47e-29 Length: 623
 Score: 43.00 Matches: 43
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 18.70% Indels: 0
 DB: 14 Gaps: 0
 US-09-938-391-2 (1-230) x CB424313 (1-623)
 QY 116 ProValValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGly 135
 |||||
 Db 570 CCGCTGCTCAACCTCAGGACGAGGTGCTGTTCTTCTAGCTGGAGGCGCTTGTCTCAGGC 511
 QY 136 SerGluGlyGlnLeuLysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeu 155
 |||||
 Db 510 TCCGAGGGCCAGCTGAAGCCGCGCCGCGCATCTTCTCTTCGACGCGCAGAGATGTCCTT 451
 QY 156 GlnHisPro 158
 |||||
 Db 450 CAGCATCCC 442
RESULT 6
 CB444165/c
 LOCUS CB444165 657 bp mRNA linear EST 25-MAR-2003
 DEFINITION 695295 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.
 ACCESSION CB444165
 VERSION CB444165.1 GI:29233914
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 657)
REFERENCE Smith,T.P.L., Roberts,A.J., Echtenkamp,S.E., Chitko-McKown,C.G.,
AUTHORS Wray,J.E. and Keele,J.W.
TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL Unpublished
COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366

Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called with phred v0.020425.c and
 trimmed with the aid of the trim_alt option. Vector identified with
 cross_match v0.990329.
 Plate: FQY8051 row: I column: 11
 Seq primer: TAGAAGGCACAGTCGAGG.
FEATURES Location/Qualifiers
 source
 1..657
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 6BOV"
 /note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
 Library made with RNA pooled from multiple tissues
 including liver, lung, hypothalamus, pituitary, and
 placenta/endometrium."
BASE COUNT 107 a 217 c 219 g 113 t
ORIGIN
 Alignment Scores:
 Pred. No.: 1.56e-29 Length: 657
 Score: 43.00 Matches: 43
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 18.70% Indels: 0
 DB: 14 Gaps: 0
 US-09-938-391-2 (1-230) x CB444165 (1-657)
 QY 116 ProValValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGly 135
 |||||
 Db 529 CCGCTGCTCAACCTCAGGACGAGGTGCTGTTCTTCTAGCTGGAGGCGCTTGTCTCAGGC 470
 QY 136 SerGluGlyGlnLeuLysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeu 155
 |||||
 Db 469 TCCGAGGGCCAGCTGAAGCCGCGCCGCGCATCTTCTCTTCGACGCGCAGAGATGTCCTT 410
 QY 156 GlnHisPro 158
 |||||
 Db 409 CAGCATCCC 401
RESULT 7
 CB443805
 LOCUS 694911 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
 DEFINITION CB443805
 ACCESSION CB443805
 VERSION CB443805.1 GI:29233191
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 709)
REFERENCE Smith,T.P.L., Roberts,A.J., Echtenkamp,S.E., Chitko-McKown,C.G.,
AUTHORS Wray,J.E. and Keele,J.W.
TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL Unpublished
COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called with phred v0.020425.c and
 trimmed with the aid of the trim_alt option. Vector identified with
 cross_match v0.990329.
 Plate: FQY8051 row: I column: 11
 Seq primer: GTAATACGACTCATTAGGG.
FEATURES Location/Qualifiers

```

source
1. .709
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."
BASE COUNT 96 a 288 c 218 g 107 t
ORIGIN

Alignment Scores:
Pred. No.: 1.68e-29 Length: 709
Score: 4300 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.70% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-2 (1-230) x CB443805 (1-709)

QY 70 ArgGlyIleArgGlyAlaAspPheGlnCysPheGlnAlaArgAlaAlaGlyLeuAla 89
Db 559 CGCGCATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 618
QY 90 GlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIleValArgArg 109
Db 619 GGCACGCTCCGCGCGGTCCTGCTCCTCGCGGTTGCAGACCTGTACAGCATCGTGGCGCG 678
QY 110 AlaAspArg 112
Db 679 GCCGACCGT 687

RESULT 8
BF601253 387 bp mRNA linear EST 25-APR-2001
LOCUS
DEFINITION 266182 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BF601253
VERSION BF601253.1 GI:11698475
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 387)
Smith,T.P.L., Freking,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-Mckown,C.G.,
Perlea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keele,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 526-630 (2001)
21180013
11282978
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACGTCAGCAGC
Plate: 40 row: H column: 7
Seq primer: ATTAGGTGACACTATAG.

```

```

FEATURES
source
1. .387
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 3BOV"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendinosus muscle, and fetal
longissimus muscle."
BASE COUNT 52 a 140 c 131 g 64 t
ORIGIN

Alignment Scores:
Pred. No.: 3.2e-24 Length: 387
Score: 3700 Matches: 37
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.09% Indels: 0
DB: 10 Gaps: 0

US-09-938-391-2 (1-230) x BF601253 (1-387)

QY 76 AspPheGlnCysPheGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPhe 95
Db 1 GACTTTCAGTCTCCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGTC 60
QY 96 LeuSerSerArgLeuGlnAspLeuTyrSerIleValArgArgAlaAspArg 112
Db 61 CTGTCTCGCGGTTCCAGGACCTGTACAGCATCGTGGCGCGCGCGCGCGCGCGT 111

RESULT 9
BG823096 483 bp mRNA linear EST 22-MAY-2001
LOCUS
DEFINITION 602726433F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4865936 5',
mRNA sequence.
ACCESSION BG823096
VERSION BG823096.1 GI:14170683
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 483)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1732 row: a column: 09
High quality sequence stop: 479.
Location/Qualifiers
1. .483
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4865936"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_15"
/note="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average

```

insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"

BASE COUNT 69 a 200 c 146 g 68 t
ORIGIN

Alignment Scores:
Pred. No.: 1-71e-19 Length: 483
Score: 32.00 Matches: 45
Percent Similarity: 97.83% Conservative: 0
Best Local Similarity: 97.83% Mismatches: 0
Query Match: 13.91% Indels: 1
DB: 12 Gaps: 0

US-09-938-391-2 (1-230) x BG823096 (1-483)

QY 67 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaAla 86
DB 270 GCGCGCATCGGGGCGTCCGGGGCGGCGACTTCAGTGTTCAGCAGCGGGCGGCGCT 329

QY 87 -GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSeril 106
DB 330 GGGGCTGGCGGCGACCTTCGCGCTTCTCTCTCGCGCTGCGGCGTGTACAGCAT 389

QY 106 eValArgArgAlaAsp 111
DB 390 CGTGGCGGCGGCCAC 405

RESULT 10
LOCUS

DEFINITION
Bm488074 577 bp mRNA linear EST 07-FEB-2002
Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
clone pgm2n.pk006.h17 5' similar to gb|AAC33294.2 (AF083440)
collagen XVIII [Gallus gallus], mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Bm488074.1 GI:18609005
Gallus gallus (chicken)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

1 (bases 1 to 577)

Cogburn,L.A. and Monsonego-Ornan,E.

ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and
Epiphyseal Growth Plate cDNA library, USDA/IFAFS Animal Genome
Project

Unpublished

Contact: Larry A. Cogburn

University of Delaware

Townsend Hall, Newark, DE 19717, USA

Tel: 302-831-1335

Fax: 302-831-2822

Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES
source

1. .577
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Commercial broiler and Ottawa Res. Centre
Strains 90 & 21"
/db_xref="taxon:9031"
/clone="pgm2n.pk006.h17"
/tissue_type="Breast muscle, leg muscle and epiphyseal
growth plate"
/dev_stage="Breast, leg, embryo(d19); post-hatch(1d,1.3,5,7,9
11 weeks); growth plate(1d,7d,14d post-hatch)"
/lab_host="E. coli BMDH108"
/clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
and Epiphyseal Growth Plate cDNA library (pgm2n)"
/note="Vector: pCMVSPORT6; Library made from equivalent

pools of total RNA isolated from each tissue (embryonic
muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
plate 33.3% of the final RNA pool). Single pass sequencing
from 5'-end"

BASE COUNT 91 a 207 c 169 g 101 t 9 others
ORIGIN

Alignment Scores:
Pred. No.: 8-68e-15 Length: 577
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.74% Indels: 0
DB: 12 Gaps: 0

US-09-938-391-2 (1-230) x Bm488074 (1-577)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSeril 106
DB 259 GGGCTGGCGGTACCTTCGTGCTTCTCTCTCCGCTGCGGACCTGTACAGCATC 318

QY 107 ValArgArgAlaAspArgThr 113
DB 319 GTGCGCAGGGCGCGCCACT 339

RESULT 11

Bm459935

LOCUS

DEFINITION

Bm459935 614 bp mRNA linear EST 29-NOV-2002

603367327F1 CSEQRBN19 Gallus gallus cDNA clone CHEST286f8 5', mRNA

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Gallus gallus (chicken)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 614)

Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,

Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.

A Comprehensive Collection of Chicken cDNAs

Curr. Biol. 12 (22), 1965-1969 (2002)

22335534

PUBMED

12445392

Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology (UMIST

)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1. .614

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Layer"

/db_xref="taxon:9031"

/clone="CHEST268f8"

/sex="Female"

/dev_stage="adult"

/lab_host="DH108"

/clone_lib="CSEQRBN19"

/note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:

EcoRI; Site_2: NotI; This normalized library was

constructed from 1 million independent clones. cDNA

synthesis was initiated using an oligo(dT) primer, using

methylated C in the first strand reaction, double-stranded cDNA

Following this first strand reaction, double-stranded cDNA

was bluntended, ligated to NotI adapters, digested with EcoRI

, size-selected, and cloned into the NotI and EcoRI

compatible sites of a custom modified MCS of the

pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

BASE COUNT 106 a 212 c 179 g 117 t

ORIGIN

Alignment Scores:
Pred. No.: 9,25e-15 Length: 614
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.74% Indels: 0
DB: 13 Gaps: 0

US-09-938-391-2 (1-230) x BU459935 (1-614)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
|||
Db 183 GGGCTGGCGGTACCTTCGTCCTCTCTCCCGCTGCAGGACCTGTACAGCATC 242
|||
QY 107 ValArgArgAlaAspArgThr 113
|||
Db 243 GTGCGCAGGCGCGACCGCACT 263
|||

RESULT 12

LOCUS BU352506

DEFINITION 603527982F1 CSEQCHN69 Gallus gallus cDNA clone CHEST477d22 5', mRNA sequence.

ACCESSION BU352506

VERSION BU352506.1 GI:25860507

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

1 (bases 1 to 652)

Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken CDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)

22335534

12445392

Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology (UMIST)

COMMENT

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

Location/Qualifiers

1..652

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Compton Line 151"

/db_xref="taxon:9031"

/clone="CHEST477d22"

/sex="Female"

/tissue_type="cerebellum"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="CSEQCHN69"

/note="Organ: brain; Vector: pBluescript II KS(+); Site 1:

ECORI; Site 2: NotI; This normalized library was

constructed from 1 million independent clones. cDNA

synthesis was initiated using an oligo(dT) primer, using

methylated C in the first strand synthesis reaction.

Following this first strand reaction, double-stranded cDNA

was blunted, ligated to NotI adapters, digested with EcoRI

, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

BASE COUNT 116 a 232 c 185 g 119 t

ORIGIN

Alignment Scores:

Pred. No.: 9,85e-15 Length: 652
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.74% Indels: 0
DB: 13 Gaps: 0

US-09-938-391-2 (1-230) x BU352506 (1-652)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
|||

Db 221 GGGCTGGCGGTACCTTCGTCCTCTCTCCCGCTGCAGGACCTGTACAGCATC 280
|||

QY 107 ValArgArgAlaAspArgThr 113
|||

Db 281 GTGCGCAGGCGCGACCGCACT 301
|||

RESULT 13

LOCUS CD215096

DEFINITION

CD215096 678 bp mRNA linear EST 20-MAY-2003
pgm2n.pk013.n20 Normalized chicken muscle cDNA library (pgm2n)
Gallus gallus cDNA clone pgm2n.pk013.n20 5' similar to
gb|AAC3294.2 (AF083440) collagen XVIII [Gallus gallus], mRNA
sequence.

ACCESSION CD215096

VERSION CD215096.1 GI:30954764

KEYWORDS EST.

SOURCE

ORGANISM

Gallus gallus (chicken)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 678)

Cogburn, L.A. and Monsonego-Ornan, E.

Chicken ESTs from muscle

Unpublished

Contact: Larry A. Cogburn

University of Delaware

Townsend Hall, Newark, DE 19717, USA

Tel: 302-831-1335

Fax: 302-831-2822

Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES

source

Location/Qualifiers

1..678

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Commercial broiler chickens, Ottawa Research Ctr

Strains 90 & 21"

/db_xref="taxon:9031"

/clone="pgm2n.pk013.n20"

/sex="Male and Female"

/tissue_type="Breast muscle, leg muscle and epiphyseal

growth plate"

/dev_stage="Breast, leg: Embryo (dl9) post-hatch (dl1,1,3,5,7,9

,11 weeks); growth plate (dl,7d,14d post-hatch)"

/lab_host="E. Coli EMDH10B"

/clone_lib="Normalized chicken muscle cDNA library (pgm2n

)"

/note="Vector: PCMVSPORT6; Library made from equivalent

pools of total RNA isolated from each tissue (embryonic

muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth

plate 33.3% of the final RNA pool). Single pass sequencing

adapted from Soares et al., PNAS (1994) 91: 9228-9232 and
Bonaldo et al., Genome Research 6 (1996): 791, except that
a significantly longer reannealing hybridization was
used "

BASE COUNT 125 a 253 c 202 g 126 t
ORIGIN

Alignment Scores:
Pred. No.: 1.07e-14 Length: 706
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.74% Indels: 0
DB: 13 Gaps: 0

US-09-938-391-2 (1-230) x BU439577 (1-706)

QY 87 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 106
DB 279 GGGCTGGCCGGTACCTTCGGTGCCTTCCTCTCTCCCGCTGCAGGACCTGTACAGCATC 338
QY 107 ValArgArgAlaAspArgThr 113
DB 339 GTGGCAGGGCCGACCGCACT 359

Search completed: August 18, 2003, 03:41:27.
Job time : 2045.89 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 15:15:38 ; Search time 2786.23 Seconds
(without alignments)
8148.953 Million cell updates/sec

Title: US-09-938-391-3
Perfect score: 555
Sequence: 1 cacaccaccagacttcca.....tgacctcttctccaagtag 555

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 2889711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pin:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	555	100.0	555	6	AX399631	Sequence
2	555	100.0	829	6	AX399629	Sequence
3	26	4.7	534	6	AR193166	Sequence
4	26	4.7	537	6	AX100092	Sequence
5	26	4.7	546	6	AR193165	Sequence
6	26	4.7	549	6	AX100086	Sequence
7	26	4.7	552	6	AX395662	Sequence
8	26	4.7	552	6	AX473835	Sequence
9	26	4.7	552	6	BD081407	Fused pro
10	26	4.7	555	9	AF184060	Homo sapi
11	26	4.7	564	9	AF416592	Homo sapi
12	26	4.7	650	6	AX370853	Sequence
13	26	4.7	786	9	AF282883	Homo sapi
14	26	4.7	900	6	AX370851	Sequence
15	26	4.7	2364	9	AK098216	Homo sapi
16	26	4.7	2837	9	BC033715	Homo sapi
17	26	4.7	3394	6	AX409531	Sequence
18	26	4.7	3394	6	IS1045	Sequence 4
19	26	4.7	3394	9	HUMCOL18AX	L22548 Human colla
20	26	4.7	5408	9	AF018082	Homo sapi
21	26	4.7	5929	9	AF018081	Homo sapi
22	26	4.7	131563	9	HSE12310	Homo sapi
23	26	4.7	223201	9	HS53110	Homo sapi
24	26	4.7	340000	9	HS21C102	Homo sapi
25	22	4.0	515	10	RNO236873	Rattus no
26	22	4.0	851	10	AF189709	Rattus no
27	22	4.0	5279	5	AF083440	Gallus ga
28	22	4.0	161250	8	AC079634	Genomic S
29	22	4.0	238671	2	AC125296	Rattus no
30	22	4.0	300029	8	AE017080	Oryza sat
31	20	3.6	306	9	AF333247	Homo sapi
32	20	3.6	307	9	AF333246	Homo sapi
33	20	3.6	919	10	F261109S22	Mus muscu
34	20	3.6	5126	10	BC052428	Mus muscu
35	20	3.6	5176	10	AF011450	Mus muscu
36	20	3.6	89212	9	AC016679	Homo sapi
37	20	3.6	93835	2	AC021904	Homo sapi
38	20	3.6	110000	2	LMFLCHR34_07	Continuation (8 of
39	20	3.6	134841	9	AC018840	Homo sapi
40	20	3.6	144663	2	AP004995	Oryza sat
41	20	3.6	155568	8	NCB19A17	Neurospor
42	20	3.6	155825	9	AC066584	Homo sapi
43	20	3.6	156218	2	AP004994	Homo sapi
44	20	3.6	157034	10	AL772232	Mouse DNA
45	20	3.6	168812	9	AC011327	Homo sapi

ALIGNMENTS

RESULT 1
AX399631
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1
Sheppard, M.G. and Tong, X.
AUTHORS
TITLE
Methods and compositions for diagnosing and treating disorders
involving angiogenesis

AX399631
Sequence 3 from Patent EP1191036.
AX399631.1
GI:21335410
555 bp
DNA
linear
PAT 06-JUN-2002


```

Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGTCTCCAGCAGGCGG 113
Db 76 GACTTCCAGTGTCTCCAGCAGGCGG 101

RESULT 4
AX100092
LOCUS AX100092 537 bp DNA linear PAT 02-APR-2001
DEFINITION Sequence 10 from Patent WO0119989.
ACCESSION AX100092
VERSION AX100092.1 GI:13539063
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Liang, H., Sim, K.L., Chang-Murad, A., Zhou, X., Madsen, J.,
Boerner, R.J., Bermejo, L.L., Mistry, F.R., Shepard, S.R. and
Schrimsher, J.L.
TITLE Method of producing and purifying endostatin?tm protein
JOURNAL Patent: WO 0119989-A 10 22-MAR-2001;
EntreMed, Inc. (US)
FEATURES
source
1..537
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 79 a 189 c 176 g 93 t
ORIGIN

Query Match 4.7%; Score 26; DB 6; Length 537;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGTCTCCAGCAGGCGG 113
Db 76 GACTTCCAGTGTCTCCAGCAGGCGG 101

RESULT 5
AR193165
LOCUS AR193165 546 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 4 from patent US 6346510.
ACCESSION AR193165
VERSION AR193165.1 GI:20239130
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 546)
AUTHORS O'Reilly, M.S. and Folkman, M. Judah.
TITLE Therapeutic antiangiogenic endostatin compositions
JOURNAL Patent: US 6346510-A 4 12-FEB-2002;
EntreMed, Inc. (US)
FEATURES
source
1..546
/organism="unknown"
BASE COUNT 80 a 196 c 177 g 93 t
ORIGIN

Query Match 4.7%; Score 26; DB 6; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGTCTCCAGCAGGCGG 113
Db 88 GACTTCCAGTGTCTCCAGCAGGCGG 113

RESULT 6

Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGTCTCCAGCAGGCGG 113
Db 88 GACTTCCAGTGTCTCCAGCAGGCGG 113

RESULT 7
AX395662
LOCUS AX395662 552 bp DNA linear PAT 18-MAY-2002
DEFINITION Sequence 30 from Patent WO0193897.
ACCESSION AX395662
VERSION AX395662.1 GI:21066471
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Sim, K.L. and Macdonald, N.J.
TITLE Angiostatin and endostatin binding proteins and methods of use
JOURNAL Patent: WO 0193897-A 30 13-DEC-2001;
EntreMed, Inc. (US)
FEATURES
source
1..552
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 83 a 196 c 179 g 94 t
ORIGIN

Query Match 4.7%; Score 26; DB 6; Length 552;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGTCTCCAGCAGGCGG 113
Db 88 GACTTCCAGTGTCTCCAGCAGGCGG 113

RESULT 8
AX473835
LOCUS AX473835 552 bp DNA linear PAT 09-AUG-2002
DEFINITION Sequence 2 from Patent WO0230982.
ACCESSION AX473835
VERSION AX473835.1 GI:22208005

```

```

AX100086
LOCUS AX100086 549 bp DNA linear PAT 02-APR-2001
DEFINITION Sequence 4 from Patent WO0119989.
ACCESSION AX100086
VERSION AX100086.1 GI:13539061
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Liang, H., Sim, K.L., Chang-Murad, A., Zhou, X., Madsen, J.,
Boerner, R.J., Bermejo, L.L., Mistry, F.R., Shepard, S.R. and
Schrimsher, J.L.
TITLE Method of producing and purifying endostatin?tm protein
JOURNAL Patent: WO 0119989-A 4 22-MAR-2001;
EntreMed, Inc. (US)
FEATURES
source
1..549
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 82 a 196 c 178 g 93 t
ORIGIN

Query Match 4.7%; Score 26; DB 6; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGTCTCCAGCAGGCGG 113
Db 88 GACTTCCAGTGTCTCCAGCAGGCGG 113

RESULT 7
AX395662
LOCUS AX395662 552 bp DNA linear PAT 18-MAY-2002
DEFINITION Sequence 30 from Patent WO0193897.
ACCESSION AX395662
VERSION AX395662.1 GI:21066471
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Sim, K.L. and Macdonald, N.J.
TITLE Angiostatin and endostatin binding proteins and methods of use
JOURNAL Patent: WO 0193897-A 30 13-DEC-2001;
EntreMed, Inc. (US)
FEATURES
source
1..552
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 83 a 196 c 179 g 94 t
ORIGIN

Query Match 4.7%; Score 26; DB 6; Length 552;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGTCTCCAGCAGGCGG 113
Db 88 GACTTCCAGTGTCTCCAGCAGGCGG 113

RESULT 8
AX473835
LOCUS AX473835 552 bp DNA linear PAT 09-AUG-2002
DEFINITION Sequence 2 from Patent WO0230982.
ACCESSION AX473835
VERSION AX473835.1 GI:22208005

```



```
PUBMED 11517600
REFERENCE 2 (bases 1 to 564)
AUTHORS Feng, Y., Wu, Y., Zhu, X., Liu, C. X. and Ma, Q. J.
TITLE Endostatin contributes to maintain cartilage homeostasis via
        promotion of the anabolic program of chondrocytes
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 564)
AUTHORS Feng, Y., Cui, L. B. and Ma, Q. J.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-2001) Genetic Engineering, Beijing Institute of
        Biotechnology, Taiping Road, Beijing 100850, P.R. China
FEATURES             Location/Qualifiers
     source          1..564
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="21"
     CDS              41..564
                     /note="similar to endostatin; can inhibit endothelial cell
                     proliferation"
                     /codon_start=1
                     /product="collagen XVIII"
                     /protein_id="AAL37720.1"
                     /db_xref="GI:17282898"
                     /translation="SPPAHSHRDFOFVLHLVALNSPLSGMRGIRGADFQCFQARAV
                     GLAGTFRLQSLQDLYSIVRRADRAAVPIVNLKDELFFPSWEALFSGSEGPKLPGA
                     RIFPDGDLRLHPTWPKSVVHGDSPNGRLTESYCETWRTAPSATGQASSLLGGR
                     LLGSAASHAHYIVLCIENSFWTASK"
     misc_feature     41..561
                     /note="Region: NC1 domain"
BASE COUNT   86 a 203 c 180 g 95 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
25
```

Best Local Similarity 100.0%; Pred. NO. 0.13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGCTTCCAGCAGGCGG 113
|||||
DB 424 GACTTCCAGTGCTTCCAGCAGGCGG 449

RESULT 15

AK098216 2364 bp mRNA linear PRI 15-JUL-2002
LOCUS Homo sapiens cDNA FLJ40897 fis, clone UTERU2002964, highly similar
DEFINITION to Homo sapiens type XVIII collagen (COL18A1) mRNA.

AK098216
ACCESSION AK098216.1 GI:21758185
VERSION
KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens
ORGANISM

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1
Oshima, A., Takahashi-Fujii, A., Tanase, T., Inose, N., Takeuchi, K.,
Arita, M., Musashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuna, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B.,
Suzuki, Y., Sugano, S., Negahari, K., Masuho, Y., Nagai, K. and
Isogai, T.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2364)

AUTHORS Isogai, T. and Yamamoto, J.

TITLE Direct Submission

JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5' - & 3' - end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

FEATURES

source Location/Qualifiers
1..2364
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UTERU2002964"
/tissue_type="uterus"
/clone_lib="UTERU2"
/note="cloning vector: pME18SFL3"

BASE COUNT 420 a 768 c 776 g 400 t
ORIGIN

Query Match 4.7%; Score 26; DB 9; Length 2364;

Best Local Similarity 100.0%; Pred. NO. 0.12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGCTTCCAGCAGGCGG 113
|||||
DB 1731 GACTTCCAGTGCTTCCAGCAGGCGG 1756

Search completed: August 17, 2003, 18:39:19
Job time : 2788.23 secs

QY 88 GACTTCCAGTGCTTCCAGCAGGCGG 113
Db 88 GACTTCCAGTGCTTCCAGCAGGCGG 113

RESULT 3
US-09-206-059-30
; Sequence 30, Application US/09206059
; Patent No. 6201104
; GENERAL INFORMATION:
; APPLICANT: Macdonald, Nicholas
; APPLICANT: Sim, Kim Lee
; TITLE OF INVENTION: Angiogenesis-Inhibiting Protein Binding Peptides and
; FILE REFERENCE: 05213-0370
; CURRENT APPLICATION NUMBER: US/09/206,059
; CURRENT FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-206-059-30

Query Match 4.7%; Score 26; DB 3; Length 552;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGCTTCCAGCAGGCGG 113
Db 88 GACTTCCAGTGCTTCCAGCAGGCGG 113

RESULT 4
US-08-159-784-4
; Sequence 4, Application US/08159784
; Patent No. 5643783
; GENERAL INFORMATION:
; APPLICANT: Bjorn R. Olsen
; TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: Wordperfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,784
; FILING DATE: December 1, 1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: John F. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00246/170001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3394

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-159-784-4

Query Match 4.7%; Score 26; DB 1; Length 3394;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGCTTCCAGCAGGCGG 113
Db 1591 GACTTCCAGTGCTTCCAGCAGGCGG 1616

RESULT 5
US-09-313-294A-2718
; Sequence 2718, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Ialguidi, Raghnath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 2718
; LENGTH: 273
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700553162H1
; NAME/KEY: unsure
; LOCATION: 88, 189, 238, 243, 246, 255, 260
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-2718

Query Match 3.2%; Score 18; DB 4; Length 273;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 428 CGCGCGCCACCGCGCAGG 445
Db 98 CGCGCGCCACCGCGCAGG 115

RESULT 6
US-08-956-307B-10
; Sequence 10, Application US/08956307B
; Patent No. 6090911
; GENERAL INFORMATION:
; APPLICANT: Petka, Wendy A.
; APPLICANT: Tirrell, David A.
; APPLICANT: Kevin P. McGrath
; TITLE OF INVENTION: REVERSIBLE HYDROGELS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,307B

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 16:25:28 ; Search time 56.5426 Seconds
(without alignments)
4332.442 Million cell updates/sec

Title: US-09-938-391-3

Perfect score: 555

Sequence: 1 cacaccaccaggactcca.....tgacctcttccaagtag 555

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/prodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/prodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	4.7	534	4	US-09-315-689-6
2	26	4.7	546	4	US-09-315-689-4
3	26	4.7	552	3	US-09-206-059-30
4	26	4.7	3394	1	US-08-159-784-4
5	18	3.2	273	4	US-09-313-294A-2718
6	18	3.2	288	3	US-08-956-307B-10
7	18	3.2	292	3	US-08-956-307B-9
8	18	3.2	558	3	US-09-449-293-3
9	18	3.2	558	3	US-09-775-325-3
10	18	3.2	565	3	US-08-985-526-37
11	18	3.2	573	4	US-09-561-500-12
12	18	3.2	573	4	US-09-561-108-12
13	18	3.2	573	4	US-09-561-108-12
14	18	3.2	573	4	US-09-561-499-12
15	18	3.2	774	3	US-08-956-307B-12
16	18	3.2	774	3	US-08-956-307B-11
17	18	3.2	1577	4	US-09-252-991A-16492
18	18	3.2	1785	4	US-09-252-991A-16007
19	18	3.2	2028	4	US-09-252-991A-16110
20	18	3.2	4031	1	US-08-159-784-1
21	18	3.2	4403765	3	US-09-103-840A-2
22	18	3.2	4403765	3	US-09-103-840A-2
23	18	3.2	4411529	3	US-09-103-840A-1
24	18	3.2	4411529	3	US-09-103-840A-1
25	17	3.1	522	4	US-09-252-991A-16260
26	17	3.1	630	4	US-09-252-991A-5585
27	17	3.1	690	4	US-09-252-991A-2533

28	17	3.1	753	4	US-09-252-991A-5613	Sequence 5613, Ap
C 29	17	3.1	870	4	US-09-252-991A-2385	Sequence 2385, Ap
C 30	17	3.1	918	4	US-09-252-991A-5525	Sequence 5525, Ap
31	17	3.1	951	4	US-09-252-991A-16122	Sequence 16122, A
32	17	3.1	1065	4	US-09-252-991A-12208	Sequence 12208, A
C 33	17	3.1	1326	4	US-09-252-991A-2299	Sequence 2299, Ap
C 34	17	3.1	1422	4	US-09-252-991A-12274	Sequence 12274, A
C 35	17	3.1	1607	3	US-08-753-007A-3	Sequence 3, Appli
C 36	17	3.1	1607	3	US-09-398-496-3	Sequence 3, Appli
C 37	17	3.1	1722	4	US-09-252-991A-15090	Sequence 15090, A
C 38	17	3.1	1803	4	US-09-252-991A-12345	Sequence 12345, A
C 39	17	3.1	2286	4	US-09-073-009-9	Sequence 9, Appli
C 40	17	3.1	2373	4	US-09-252-991A-16481	Sequence 16481, A
C 41	17	3.1	2467	3	US-08-753-007A-1	Sequence 1, Appli
C 42	17	3.1	2467	3	US-09-398-496-1	Sequence 1, Appli
C 43	17	3.1	2608	4	US-09-904-615-16	Sequence 16, Appli
C 44	17	3.1	3107	2	US-08-813-940-3	Sequence 3, Appli
C 45	17	3.1	3441	2	US-08-525-864A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-315-689-6
; Sequence 6, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Folkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-315-689-6

Query Match 4.7%; Score 26; DB 4; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.017; 0; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGCTTCCAGCAGCGCG 113
DB 76 GACTTCCAGTGCTTCCAGCAGCGCG 101

RESULT 2
US-09-315-689-4
; Sequence 4, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Folkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-315-689-4

Query Match 4.7%; Score 26; DB 4; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.017; 0; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 14:11:23 ; Search time 214.14 Seconds
(without alignments)
6996.296 Million cell updates/sec

Title: US-09-938-391-3
Perfect score: 555
Sequence: 1 cacaccaccaggacttcca.....tgacctctcttccaagtag 555

Scoring table: OLIGO_NUC =
Gapop_60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	555	100.0	555	24	Canine endostatin
2	555	100.0	829	24	Canine pro-endosta
3	529	95.3	552	21	Canine angiogenesi
4	29	5.2	552	21	Human endostatin c
5	26	4.7	534	21	Alternate human en
6	26	4.7	537	22	Human gene fragmen
7	26	4.7	546	21	Human endostatin g
8	26	4.7	549	20	Human endostatin D

9	26	4.7	549	21	AAA29884	Human angiogenesis
10	26	4.7	549	21	AAZ51291	Human angiogenesis
11	26	4.7	549	22	AAZ00867	Human gene fragmen
12	26	4.7	549	25	ABA00774	Human endostatin c
13	26	4.7	551	24	ABQ81193	Human endostatin c
14	26	4.7	552	20	AAZ35375	SEQ ID 50 of WO991
15	26	4.7	552	21	AAZ62023	Nucleotide sequenc
16	26	4.7	552	21	AAA68203	Human endostatin c
17	26	4.7	552	22	AAZ88289	Human endostatin c
18	26	4.7	552	24	ABK50685	cDNA encoding huma
19	26	4.7	555	22	AAI66529	Human vascular end
20	26	4.7	558	24	ABA93261	Human endostatin c
21	26	4.7	563	20	AAZ08750	Human endostatin c
22	26	4.7	574	24	ABV94691	Human pancreatic c
23	26	4.7	641	22	AAH79104	Human endostatin e
24	26	4.7	641	24	AAI55454	Specific tumour ce
25	26	4.7	641	24	AAI44000	Reproductive recom
26	26	4.7	641	24	ABQ76079	Anticancer gene-as
27	26	4.7	650	24	ABK09978	Synthetic plasmid
28	26	4.7	816	21	AAA64013	DNA encoding a hum
29	26	4.7	900	24	ABK09977	Synthetic plasmid
30	26	4.7	968	21	AAZ62025	Nucleotide sequenc
31	26	4.7	1564	24	ABQ76740	DNA encoding human
32	26	4.7	3394	18	AAZ84484	Human alpha-1 coll
33	26	4.7	3394	20	AAZ78379	Human alpha1 (XVII
34	26	4.7	3394	24	ABN95680	Gene #2178 used to
35	26	4.7	4551	24	ABV94763	Human pancreatic c
36	26	4.7	4551	24	ABN85301	Human collagen XVI
37	26	4.7	4875	24	ABQ54955	Human ovarian anti
38	26	4.7	5408	20	AAZ77720	Human collagen 18
39	26	4.7	6462	21	AAA64014	Nucleotide sequenc
C 40	26	4.7	37664	22	AAK83781	Human immune/haema
C 41	26	4.7	61710	22	AAK83782	Human immune/haema
C 42	24	4.3	34	24	AAI46067	Canine endostatin
C 43	24	4.3	34	24	AAI46071	Canine collagen XV
C 44	24	4.1	33	24	AAI46066	Canine endostatin
45	22	4.0	32	24	AAI46070	Canine collagen XV

ALIGNMENTS

RESULT 1
AAL46063
ID AAL46063 standard; cDNA; 555 BP.
XX

AC AAL46063;

DT 19-JUL-2002 (first entry)

XX Canine endostatin coding sequence.

XX Dog; pro-endostatin; endostatin; angiogenesis; cancer; gene therapy;
KW psoriasis; rheumatoid arthritis; retinopathy; macular degeneration;
KW corneal graft rejection; neovascular glaucoma; retrolental fibroplasia;
KW rubeosis; Osler-Weber Syndrome; myocardial angiogenesis;
KW plaque neovascularisation; telangiectasia; haemophilic joints;
KW angiofibroma; wound granulation; coronary collateral;
KW cerebral collateral; arteriovenous malformation;
KW ischaemic limb angiogenesis; diabetic neovascularisation; fracture;
KW cytostatic; antiarthritic; antipsoriatic; antidiabetic; ophthalmological;
KW gynaecological; gene; ss.

OS Canis familiaris.

XX Key Location/Qualifiers
CDS 1..555
FT /*tag a
FT /product= "endostatin"
FT /partial
FT /note= "no start codon"

PN EP1191036-A2.

XX PD 27-MAR-2002.
 XX PF 24-AUG-2001; 2001EP-0307224.
 XX PR 25-AUG-2000; 2000US-227924P.
 XX PA (PFIZ) PFIZER PROD INC.
 XX PI Sheppard MG, Tong X;
 XX DR WPI; 2002-354068/39.
 XX DR P-PSDB; AA017430.
 XX PT An isolated nucleic acid molecule for the treatment of
 PT angiogenesis-related disorder, such as cancers or diabetic retinopathy,
 PT encodes an endostatin protein -
 XX Claim 2; Fig 4; 56pp; English.
 XX The present invention provides the protein and coding sequences of canine
 CC pro-endostatin and endostatin. The sequences can be used in the treatment
 CC and diagnosis of angiogenesis related disorders, including cancer,
 CC rheumatoid arthritis, psoriasis, retinopathy, macular degeneration,
 CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasia,
 CC rubecosis, Osler-Webber Syndrome, myocardial angiogenesis, plaque
 CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,
 CC wound granulation, coronary collaterals, cerebral collaterals,
 CC arteriovenous malformations, ischaemic limb angiogenesis, diabetic
 CC neovascularisation, and fractures. The present sequence is the canine
 CC endostatin coding sequence.
 XX Sequence 555 BP; 78 A; 203 C; 191 G; 83 T; 0 other;
 SQ Query Match 100.0%; Score 555; DB 24; Length 555;
 Best Local Similarity 100.0%; Pred. No. 1.7e-239;
 Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 CACACCACAGGACTCCAGCTGGTCTGACCTGGTGGCCCTGACAGCCGCGAGCGG 60
 DB 1 CACACCACAGGACTTCAGCTGGTCTGACCTGGTGGCCCTGACAGCCGCGAGCGG 60
 OY 61 GGGCGCATGCGAGGACATCCCGGGAGGCGACTTCAGTGTCTTCAGCAGGCGCGCGCGG 120
 DB 61 GGGCGCATGCGAGGACATCCCGGGAGGCGACTTCAGTGTCTTCAGCAGGCGCGCGG 120
 OY 121 GGGCTGGCGGACCTTCGCGGCTTCCTGCTGCTGGGCTGAGGACCTTACAGCATC 180
 DB 121 GGGCTGGCGGACCTTCGCGGCTTCCTGCTGCTGGGCTGAGGACCTTACAGCATC 180
 OY 181 GTGCGCGCGCGGACCGGACCGGCGTCCGCTGCTGCTGAGGACGAGGTGCTTTC 240
 DB 181 GTGCGCGCGCGGACCGGACCGGCGTCCGCTGCTGCTGAGGACGAGGTGCTTTC 240
 OY 241 CCAGGTGGGAGGCTTATTTCTCGGGCTCCGAGGCGGAGTGAAGCCCGGCGCGCATC 300
 DB 241 CCAGGTGGGAGGCTTATTTCTCGGGCTCCGAGGCGGAGTGAAGCCCGGCGCGCATC 300
 OY 301 TTCTTTTTCAGCGGAGAGATGCTGACAGACACCCGCGCTGGCCCGGAGAGCGGTGG 360
 DB 301 TTCTTTTTCAGCGGAGAGATGCTGACAGACACCCGCGCTGGCCCGGAGAGCGGTGG 360
 OY 361 CAGGGTCCGACCCGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 DB 361 CAGGGTCCGACCCGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 OY 421 GAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
 DB 421 GAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
 OY 481 GAGGCGCGGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 540
 DB 481 GAGGCGCGGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 540

OY 541 TCCTTCTCCAGTAG 555
 DB 541 TCCTTCTCCAGTAG 555
 RESULT 2
 AAL46062
 ID AAL46062 standard; cDNA; 829 BP.
 XX AAL46062;
 AC AAL46062;
 XX 19-JUL-2002 (first entry)
 DT 19-JUL-2002 (first entry)
 XX Canine pro-endostatin coding sequence.
 DE Dog; pro-endostatin; endostatin; angiogenesis; cancer; gene therapy;
 XX psoriasis; rheumatoid arthritis; retinopathy; macular degeneration;
 KW corneal graft rejection; neovascular glaucoma; retrolental fibroplasia;
 KW rubecosis; Osler-Webber Syndrome; myocardial angiogenesis;
 KW plaque neovascularisation; telangiectasia; haemophilic joints;
 KW angiofibroma; wound granulation; coronary collaterals;
 KW cerebral collaterals; arteriovenous malformation;
 KW ischaemic limb angiogenesis; diabetic neovascularisation; fracture;
 KW cytostatic; antiarthritic; antipsoriatic; antidiabetic; ophthalmological;
 KW gynaecological; gene; ss.
 XX Canis familiaris.
 OS Canis familiaris.
 XX Key Location/Qualifiers
 FH CDS 1..693
 FT /*tag= a
 FT /product= "pro-endostatin"
 FT /partial
 FT /note= "no start codon"
 FT
 PN EPL191036-A2.
 XX 27-MAR-2002.
 PD 24-AUG-2001; 2001EP-0307224.
 XX 25-AUG-2000; 2000US-227924P.
 PR (PFIZ) PFIZER PROD INC.
 PA Sheppard MG, Tong X;
 PI WPI; 2002-354068/39.
 XX P-PSDB; AA017429.
 DR An isolated nucleic acid molecule for the treatment of
 PT angiogenesis-related disorder, such as cancers or diabetic retinopathy,
 PT encodes an endostatin protein -
 XX Claim 2; Fig 2; 56pp; English.
 PS The present invention provides the protein and coding sequences of canine
 CC pro-endostatin and endostatin. The sequences can be used in the treatment
 CC and diagnosis of angiogenesis related disorders, including cancer,
 CC rheumatoid arthritis, psoriasis, retinopathy, macular degeneration,
 CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasia,
 CC rubecosis, Osler-Webber Syndrome, myocardial angiogenesis, plaque
 CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,
 CC wound granulation, coronary collaterals, cerebral collaterals,
 CC arteriovenous malformations, ischaemic limb angiogenesis, diabetic
 CC neovascularisation, and fractures. The present sequence is the canine
 CC pro-endostatin coding sequence.
 XX Sequence 829 BP; 124 A; 314 C; 278 G; 113 T; 0 other;
 SQ Query Match 100.0%; Score 555; DB 24; Length 829;
 Best Local Similarity 100.0%; Pred. No. 1.7e-239;

Matches	555	Conservative	0	Mismatches	0	Indels	0	Gaps	0
Qy	1	CACACCA	CCAGGACTTCCAGCTGGTGTCTGCACCTGTGTGGCCCTGAA	CAGCCCGCAGCCG	60				
Db	139	CACACCA	CCAGGACTTCCAGCTGGTGTCTGCACCTGTGTGGCCCTGAA	CAGCCCGCAGCCG	198				
Qy	61	GGCGCAT	CGGAGCATCCGGGAGCGGACTTCCAGTGTCTCCAGCAGGCGCGCGCG	120					
Db	199	GGCGCAT	CGGAGCATCCGGGAGCGGACTTCCAGTGTCTCCAGCAGGCGCGCGCG	258					
Qy	121	GGGCTGG	CCGGACCTTCTCCGGGCCCTTCTGTCTGTCTGGCGGTCTGACGACCTCTACAGCATC	188					
Db	259	GGGCTGG	CCGGACCTTCTCCGGGCCCTTCTGTCTGTCTGGCGGTCTGACGACCTCTACAGCATC	318					
Qy	181	GTGGCC	CGCGCGACCGGACCGGGGTGCGTCTGTCAACTCAGGACGAGGTGCTCTTC	240					
Db	319	GTGGCC	CGCGCGACCGGACCGGGGTGCGTCTGTCAACTCAGGACGAGGTGCTCTTC	378					
Qy	241	CCAGCT	TGGGAGGCTTATTCTCGGGTCTCGAGGGCCAGCTGAAGCCGGGGCCCGCATC	300					
Db	379	CCAGCT	TGGGAGGCTTATTCTCGGGTCTCGAGGGCCAGCTGAAGCCGGGGCCCGCATC	438					
Qy	301	TTCTCT	TTTGACGGCAGAGATGTCTTGACGACCCCGCTTGGCCCCCGAAGAGCGTGTGG	360					
Db	439	TTCTCT	TTTGACGGCAGAGATGTCTTGACGACCCCGCTTGGCCCCCGAAGAGCGTGTGG	498					
Qy	361	CAGGCT	CCGACCTCAGCGGGGCGCGCTGTACCGACAGCTACTCGGAGACGTGCGGACG	420					
Db	499	CAGGCT	CCGACCCAGCGGGCGCGCTGTACCGACAGCTACTCGGAGACGTGCGGACG	558					
Qy	421	GAGGCC	CGGGCCACCGGCGAGGCGTCTGTCTGTGGCGGCGAGGCTGCTGGAGCAG	480					
Db	559	GAGGCC	CGGGCGGACCGGCGAGGCGTCTGTCTGTGGCGGCGAGGCTGCTGGAGCAG	618					
Qy	481	GAGGCC	CGGAGCTGCGCGCACGCTCTGTGTGTGTCTGTGCATCGAGAACAGCGTCAATACC	540					
Db	619	GAGGCC	CGGAGCTGCGCGCACGCTCTGTGTGTGTCTGTGCATCGAGAACAGCGTCAATACC	678					
Qy	541	TCCTT	TCTCCAGTAG	555					
Db	679	TCCTT	TCTCCAGTAG	693					

RESULT 3

RESULT 3
AAZ51309

AAZ51309
ID AAZ51309 standard: DNA: 552 BP.

XX
ID
AA251309XX
AC
AAZ51309.AC
XX
AAZ51309;XX
DT
06-JUN-30DT
06-JUN-20
yy

XX

DE Canine an
vv

XX
XX
XX

KW Canine; i

KW angiogene.

KW antipsori

KW vasotropi.

KW metastasi

KW. ocular an

KW myocardia

KW wound gra

XX

OS Canis fam

XX
XX

Key

FT	KEY
FT	CDS

THE
C3

二五

$$\frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

LET

FLYY

XX
DN
W03000170

PN WO2000110
yy

XX

02-MAR-2000.

25-AUG-1999; 99WO-US19329.

25-AUG-1998; 98US-0097883.

(LEXI-) LEXINGEN PHARM CORP.

Lo K, Li Y, Gillies SD;
WPI; 2000-237616/20.
P-PSDB; AAY70265.

Novel fusion protein of angiostatin or endostatin and an immunoglobulin FC region, useful for treating conditions mediated by angiogenesis, such as rheumatoid arthritis, tumors and macular degeneration -

Example 8; Pages 58-59; 68pp; English.

The patent discloses a DNA molecule encoding a fusion protein comprising a signal sequence, an immunoglobulin Fc region, and an angiogenesis inhibitor selected from angiostatin, endostatin, a plasminogen fragment having angiostatin activity, a collagen XVII fragment having endostatin activity, or combinations of them. The fusion protein (immunofusin) is used to inhibit angiogenesis and to treat diseases or conditions mediated by angiogenesis. Conditions that may be treated include solid tumours, blood born tumours, tumour metastasis, benign tumours including haemangiomas, acoustic neuromas, neurofibromas, trachomas and pyrogenic granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases e.g. diabetic retinopathy, retinopathy of prematurity, macular degeneration, corneal graft rejection, neovascular glaucoma, retrolental fibroplasia, rubeosis and Oster-Weber syndrome; myocardial angiogenesis, plaque neovascularisation, telangiectasia, haemophilic joints, angiofibroma, wound granulation, and excessive or abnormal stimulation of endothelial cells, intestinal cells, atherosclerosis, scleroderma and hypertrophic scars, i.e. keloid scars. The DNA constructs may be used in gene therapy. The present sequence is a DNA encoding canine endostatin used in the construction of immunofusin containing canine immunoglobulin Fc fragment.

Sequence 552 BP; 77 A; 204 C; 190 G; 81 T; 0 other;

CC terminal protein of human collagen XVIII. Recombinant mouse endostatin
 CC (20 mg/kg) was administered subcutaneously to mice implanted with Lewis
 CC lung carcinomas. There was tumour mass regression non-detectable levels
 CC after 12 days of therapy due to the angiogenesis inhibitory activity of
 CC endostatin. Thus the protein is useful for treatment of angiogenesis-
 CC dependent cancers. The polynucleotide and polypeptide sequences of this
 CC endostatin are useful for treating and diagnosis of tumours, ocular
 CC angiogenic diseases, Osler-Webber syndrome, myocardial angiogenesis,
 CC plaque neovascularisation, telangiectasia, haemophilic joints,
 CC angiofibroma and wound granulation, for treatment of diseases related to
 CC excessive or abnormal stimulation of endothelial cells e.g. intestinal
 CC adhesions, atherosclerosis, scleroderma. The protein may also be useful
 CC as a birth control agent by reducing or preventing uterine
 CC vascularisation. The gene for endostatin may be isolated from cells or
 CC tissue that express high levels of endostatin, eg. tumour cells, by
 CC generating cDNA from mRNA using reverse transcriptase and then amplifying
 CC the DNA sequence.

SQ Sequence 546 BP; 80 A; 196 C; 177 G; 93 T; 0 other;

Query Match 4.7%; Score 26; DB 21; Length 546;
 Best Local Similarity 100.0%; Pred. No. 0.057;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGCTTCCAGCAGCGCG 113
 |||||
 Db 88 GACTTCCAGTGCTTCCAGCAGCGCG 113

RESULT 8
 AAX77719

ID AAX77719 standard; DNA; 549 BP.

XX AC AAX77719;

XX DT 10-AUG-1999 (first entry)

XX DE Human endostatin DNA coding region fragment.

XX KW Plasminogen; human; angiostatin; endostatin; gene therapy; vector;
 KW anti-angiogenic; attenuation; cytostatic; anti-diabetic; ophthalmology;
 KW tumour growth; solid tumour; diabetic retinopathy; retina; ss.

XX OS Homo sapiens.

XX PN WO9926480-A1.

XX PD 03-JUN-1999.

XX PF 20-NOV-1998; 98WO-US24950.

XX PR 20-NOV-1997; 97US-0975424.

XX PA (GENE-) GENETIX PHARM INC.
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.

XX PI Bachelot T, Leboulch P, Pawliuk RJ;

XX WPI; 1999-357696/30.

DR P-PSDB; AAY08693.

XX Anti-angiogenic gene therapy vectors

PS Disclosure; Page 74; 83pp; English.

XX CC This invention describes a novel viral gene therapy vector comprising a
 CC nucleic acid molecule encoding an anti-angiogenic polypeptide chosen
 CC from human or murine angiostatin, human or murine endostatin and
 CC angiogenesis-inhibiting fusions and fragments, where the viral vector is
 CC sufficiently attenuated for use in human gene therapy. The products of
 CC the invention have anti-angiogenic, cytostatic, anti-diabetic and
 CC ophthalmological activity. The vector is used in gene therapy for
 CC inhibiting tumour growth in humans harbouring a solid tumour. The vector

CC expresses an anti-angiogenic polypeptide. An additional use comprises
 CC treatment of diabetic retinopathy, where the anti-angiogenic polypeptide
 CC inhibits angiogenesis in the vicinity of the retina. The vector is
 CC administered to cells ex vivo and then administered to the patient.

XX SQ Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;

Query Match 4.7%; Score 26; DB 20; Length 549;
 Best Local Similarity 100.0%; Pred. No. 0.057;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGCTTCCAGCAGCGCG 113
 |||||
 Db 88 GACTTCCAGTGCTTCCAGCAGCGCG 113

RESULT 9
 AAA29884

ID AAA29884 standard; cDNA; 549 BP.

XX AC AAA29884;

XX DT 22-AUG-2000 (first entry)

XX DE Human angiogenesis inhibiting factor 1 encoding cDNA.

XX KW Human; angiogenesis inhibiting factor 1; IAF-1; tumour; antibody;
 KW abnormal vessel disease; ss.

XX OS Homo sapiens.

XX PN CN1244536-A.

XX PD 16-FEB-2000.

XX PF 10-AUG-1998; 98CN-0117150.

XX PR 10-AUG-1998; 98CN-0117150.

XX PA (ONCO-) INST ONCOLOGY UNDER TUMOR HOSPITAL CHINE.

XX PI Yang Z, Guo W;

XX WPI; 2000-388168/34.

DR P-PSDB; AAY90771.

XX Angiogenesis inhibiting factor 1 and its derivative useful for treating
 PT tumors -

XX Example 1; Fig 5; 41pp; Chinese.

XX CC The present sequence encodes an angiogenesis inhibiting factor (1),
 CC designated IAF-1. The present invention also describes: (1) preparation
 CC of (1) and its derivative; (2) an IAF binding acceptor and its
 CC preparation; and (3) an IAF antibody. (1) is useful for preparing new
 CC biological preparations for effectively treating various tumours and
 CC abnormal-vessel diseases. The IAF antibody is preferably a polyclonal
 CC antibody, mosaic antibody, single stranded antibody and human originated
 CC antibody.

XX SQ Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;

Query Match 4.7%; Score 26; DB 21; Length 549;
 Best Local Similarity 100.0%; Pred. No. 0.057;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGCTTCCAGCAGCGCG 113
 |||||
 Db 88 GACTTCCAGTGCTTCCAGCAGCGCG 113

RESULT 10
 AAZ51291

AAZ51291 standard; cDNA; 549 BP.
AAZ51291;
06-JUN-2000 (first entry)
Human angiogenesis inhibitor, endostatin cDNA.
Human; immunoglobulin gamma Fc fragment; endostatin; immunofusin; angiogenesis; inhibitor; cytostatic; antirheumatoid; antiarthritic; antiproliferative; antidiabetic; ophthalmological; immunosuppressant; vasotropic; vulnerable; treatment; antiarteriosclerosis; tumour; metastasis; atherosclerosis; psoriasis; rheumatoid arthritis; ocular angiogenic disease; diabetic retinopathy; macular degeneration; myocardial angiogenesis; plaque neovascularisation; telangiectasia; wound granulation; keloid scar; gene therapy; ss.
Homo sapiens.
OS XX
FH Key Location/Qualifiers
FT CDS 1..549
FT /tag= a
FT /product= "Endostatin"
FT /note= "Does not include stop codon"
FT /partial
XX WO200011033-A2.
XX PD 02-MAR-2000.
XX PF 25-AUG-1999; 99WO-US19329.
XX PR 25-AUG-1998; 98US-0097883.
XX PA (LEXI-) LEXINGEN PHARM CORP.
XX PI Lo K, Li Y, Gillies SD;
XX DR WPI; 2000-237616/20.
XX DR P-PSDB; AAY70252.
XX Novel fusion protein of angiostatin or endostatin and an immunoglobulin FC region, useful for treating conditions mediated by angiogenesis, such as rheumatoid arthritis, tumors and macular degeneration -
XX Claim 12; Pages 40-41; 68pp; English.
XX The patent discloses a DNA molecule encoding a fusion protein comprising a signal sequence, an immunoglobulin FC region, and an angiogenesis inhibitor selected from angiostatin, endostatin, a plasminogen fragment having angiostatin activity, a collagen XVIII fragment having endostatin activity, or combinations of them. The fusion protein (immunofusin) is used to inhibit angiogenesis and to treat diseases or conditions mediated by angiogenesis. Conditions that may be treated include solid tumours, blood born tumours, tumour metastasis, benign tumours including haemangiomas, acoustic neuromas, neurofibromas, trachomas and pyrogenic granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases e.g. diabetic retinopathy, retinopathy of prematurity, macular degeneration, corneal graft rejection, neovascular glaucoma, retrolental fibroplasia, rubeosis and Osler-Webber syndrome, myocardial angiogenesis, plaque neovascularisation, telangiectasia, haemophilic joints, angiofibroma, wound granulation, and excessive or abnormal stimulation of endothelial cells, intestinal cells, atherosclerosis, sclerodermal and hypertrophic scars, i.e. keloid scars. The DNA constructs may be used in gene therapy. The present sequence is a cDNA encoding human endostatin used in the construction of immunofusin containing human immunoglobulin gamma (IgG) Fc fragment.
XX Note: This sequence is stated in claim 12 as being amino acid sequence of plasminogen fragment, however, the rest of the specification refers to this sequence as being nucleotide sequence of human endostatin.
XX Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;

Query Match 4.7%; Score 26; DB 21; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 88 GACTTCCAGTGTCTCCAGCAGCGCG 113
DB 88 GACTTCCAGTGTCTCCAGCAGCGCG 113
RESULT 11
AAS00867
ID AAS00867 standard; DNA; 549 BP.
XX AAS00867;
AC AC
XX 04-JUL-2001 (first entry)
DT XX
DE Human gene fragment encoding Endostatin(TM) protein.
XX KW Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;
KW blood borne tumour; leukaemia; tumour metastasis; benign tumour;
KW haemangioma; acoustic neuroma; neurofibroma; trachoma; rubeosis;
KW pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;
KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
KW retinopathy of prematurity; macular corneal graft rejection;
KW neovascular glaucoma; retrolental fibroplasia; Osler-Webber Syndrome;
KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
KW haemophilic joint; angiofibroma; wound granulation; ds.
XX KW Homo sapiens.
OS XX
FH Key Location/Qualifiers
FT CDS 1..549
FT /tag= a
FT /product= "Endostatin(TM) C-terminus minus 3"
FT /note= "Variant produced during fermentation
reaction of Pichia pastoris harbouring an expression
plasmid containing the present sequence"
FT CDS 1..543
FT /tag= b
FT /product= "Endostatin(TM) C-terminus minus 2"
FT /note= "Variant produced during fermentation
reaction of Pichia pastoris harbouring an expression
plasmid containing the present sequence"
FT CDS 1..546
FT /tag= c
FT /product= "Endostatin(TM) C-terminus minus 1"
FT /note= "Variant produced during fermentation
reaction of Pichia pastoris harbouring an expression
plasmid containing the present sequence"
FT CDS 1..549
FT /tag= d
FT /product= "Endostatin(TM)"
FT /note= "None of the above CDSs have start or stop codons"
XX WO200119989-A2.
XX PD 22-MAR-2001.
XX PF 14-SEP-2000; 2000WO-US25166.
XX PR 14-SEP-1999; 99US-0153698.
XX (ENTR-) ENTREMED INC.
XX PA Liang H, Sim KL, Chang-Murad A, Zhou X, Madsen J, Boerner RJ;
XX PI Bermejo LL, Mistry PR, Shepard SR, Schrimsher JL;
XX DR WPI; 2001-244802/25.

DR P-PSDB; AAU00896, AAU00897, AAU00898, AAU00899.
 XX Producing Endostatin protein for treating angiogenesis mediated
 PT diseases such as solid tumours, comprises recombinantly producing the
 PT protein using an expression system, and recovering and purifying the
 XX protein
 XX Claim 6; Page 29; 67pg; English.
 XX The sequence encodes Human Endostatin(TM). The new method of the
 CC invention is useful for producing, recovering and purifying Endostatin
 CC (TM) from biological sources, such as biological fluids, tissues, cells,
 CC culture media, and fermentation media. Endostatin(TM) is useful for
 CC treating angiogenesis mediated diseases such as solid tumours, blood
 CC borne tumours, leukaemias, tumour metastases, benign tumours, e.g.
 CC haemangioma, acoustic neuromas, neurofibromas, trachomas, and pyogenic
 CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases,
 CC e.g., diabetic retinopathy, retinopathy of prematurity, macular
 CC degeneration, corneal graft rejection, neovascular glaucoma, colon
 CC cancer, retrolental fibroplasia, rubecosis, Oeler-Webber Syndrome,
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,
 CC haemophilic joints, angiofibroma, and wound granulation. Endostatin(TM)
 CC is also useful for treating disease of excessive or abnormal stimulation
 CC of endothelial cells such as intestinal adhesions, atherosclerosis,
 CC scleroderma and hypertrophic scars. Higher yields of more purified, and
 CC biologically active Endostatin(TM) are obtained by the new method.
 CC Endostatin(TM) can be stored in buffers for extended periods of time, and
 CC also subjected to lyophilisation, while preserving biological activity.
 CC Centrifugation of broth from fermentation steps in production is avoided,
 CC preventing unwanted potential cellular lysis and contamination with
 CC additional proteins, pigments, enzymes and other cellular chemicals and
 CC debris.
 XX Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;
 SQ
 Query Match 4.7%; Score 26; DB 22; Length 549;
 Best Local Similarity 100.0%; Pred. No. 0.057;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 88 GACTTCCAGTGCTTCCAGCAGGCGG 113
 DB 88 GACTTCCAGTGCTTCCAGCAGGCGG 113
 RESULT 12
 ABA00774
 ID ABA00774 standard; cDNA; 549 BP.
 XX ABA00774;
 XX
 DT 18-MAR-2003 (first entry)
 XX Human endostatin coding sequence.
 DE
 XX Gene; human; plasminogen; angiostatin; neovascularisation;
 KW kringle domain; cell proliferation; viral vector;
 KW replication-defective; cancer; tumour; ss.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FT CDS 1..549
 FT /*tag= a
 FT /product= "Endostatin"
 FT /transl_except= "(pos:2..3,aa:His)"
 FT /note= "the CDS does not include a start codon"
 XX
 PN WO200288173-A2.
 XX
 XX 07-NOV-2002.
 XX
 XX 29-APR-2002; 2002WO-US13461.
 XX
 XX 30-APR-2001; 2001US-287673P.
 PR 05-APR-2002; 2002US-370634P.
 PR

(CELL-) CELL GENESYS INC.
 Chang B, Wu WW, Macarthur J, Patel S, Jooss K, Mendez M;
 WPI; 2003-129131/12.
 DR P-PSDB; AAG79753.
 XX New recombinant viral vector expressing human angiostatin useful for
 PT inhibiting angiogenesis in a mammalian subject with cancer or tumor
 XX
 XX Example 4; Page 80-82; 83pg; English.
 XX This sequence encodes endostatin. Endostatin is a 20 kD C-terminal
 CC fragment of collagen XVIII that inhibits angiogenesis. The endostatin
 CC coding sequence may be used in the recombinant viral vector of the
 CC invention for obtaining angiostatin activity. The vector comprises a
 CC promoter capable of expressing human angiostatin operably linked to a
 CC structural gene encoding one or more domains of human angiostatin.
 CC The vector, which may be a replication-defective viral vector, is useful
 CC for inhibiting angiogenesis in a mammal, especially with cancer or a
 CC tumour.
 CC Note: This sequence is given incorrectly in the sequence listing of
 CC the specification as an amino acid sequence.
 XX Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;
 SQ
 Query Match 4.7%; Score 26; DB 25; Length 549;
 Best Local Similarity 100.0%; Pred. No. 0.057;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 88 GACTTCCAGTGCTTCCAGCAGGCGG 113
 DB 88 GACTTCCAGTGCTTCCAGCAGGCGG 113
 RESULT 13
 ABO81193
 ID ABO81193 standard; cDNA; 551 BP.
 XX ABO81193;
 AC ABO81193;
 XX
 DT 05-DEC-2002 (first entry)
 XX Human endostatin coding sequence.
 DE
 XX Endostatin; human; ophthalmological; ocular neovascularisation;
 KW choroidal neovascularisation; gene therapy; gene; ss.
 KW
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FT CDS 1..551
 FT /*tag= a
 FT /partial
 FT /product= "Endostatin"
 FT /transl_except= "(pos:2..3,aa:His)"
 FT /note= "the CDS does not include a start codon"
 XX
 PN WO200267971-A2.
 XX
 XX 06-SEP-2002.
 XX
 XX 21-FEB-2002; 2002WO-US05336.
 XX
 XX 22-FEB-2001; 2001US-270787P.
 PR 04-APR-2001; 2001US-281296P.
 XX
 XX (NOVS) NOVARTIS AG.
 XX
 XX Brazzell RK, Campochiaro PA, Dixon KH;
 PI WPI; 2002-698636/75.
 DR

DR P-PSDB; ABB79901.
XX
PT Treating or preventing choroidal neovascularization comprises
PT increasing the amount of endostatin in ocular tissues of afflicted
PT individuals to a choroidal neovascularization inhibiting level -
XX
PS Claim 27; Page 39-40; 44pp; English.
XX
CC The present sequence is a partial coding sequence for human
CC endostatin. A claimed method for the treatment of ocular
CC neovascularisation, especially choroidal neovascularisation,
CC involves increasing the level of endostatin in ocular tissue,
CC especially where the endostatin is encoded by the present sequence,
CC or is its fragment, derivative or variant. The increase is
CC effected by administering a viral vector, especially an adenovirus,
CC adeno-associated virus, a retrovirus or lentivirus vector,
CC comprising an endostatin-encoding nucleic acid. Cells secreting
CC endostatin may be encapsulated and implanted within an individual.
CC The method is used when ocular neovascularisation is caused by
CC histoplasmosis, pathological myopia, angiod streaks, anterior
CC ischaemic optic neuropathy, bacterial endocarditis, Best's disease,
CC birdshot retinochoroidopathy, choroidal haemangioma, choroidal
CC naevi, choroidal nonperfusion, choroidal osteomas, choroidoma of
CC rupture, choroideraemia, chronic retinal detachment, coloboma of
CC the retina, Drusen, endogenous Candida endophthalmitis,
CC extrapapillary hamartoma of the retinal pigmented epithelium,
CC fundus flavimaculatus, idiopathic, macular hole, malignant
CC melanoma, membranoproliferative glomerulonephritis (type II),
CC metallic intraocular foreign body, morning glory disc syndrome,
CC multiple evanescent white-dot syndrome, neovascularisation of ora
CC serrata, operating microscope burn, optic nerve head pits,
CC photocoagulation, punctate inner choroidopathy, rubella,
CC sarcoidosis, serpinginous or geographic choroiditis, subretinal
CC fluid drainage, tilted disc syndrome, toxoplasma retinochoroiditis,
CC tuberculosis, Vogt-Koyangi-Harada syndrome, diabetic retinopathy,
CC non-diabetic retinopathy, brain vein occlusion, central retinal
CC vein occlusion, retinopathy in premature infants, rubeosis iridis,
CC neovascular glaucoma, perifoveal telangiectasis, sickle cell
CC retinopathy, Eale's disease, retinal vasculitis, Von Hippel
CC Lindau disease, radiation retinopathy, retinal cryoinjury,
CC retinitis pigmentosa, retinochoroidal coloboma, corneal
CC neovascularisation due to herpes simplex keratitis, corneal
CC keratoplasty, pterygia and trauma (all claimed).
XX
SQ Sequence 551 BP; 83 A; 195 C; 179 G; 94 T; 0 other;

Query Match 4.7%; Score 26; DB 24; Length 551;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGTTCAGCAGCGCG 113
Db 87 GACTTCCAGTGTTCAGCAGCGCG 112

RESULT 14
AAX35375
ID AAX35375 standard; DNA; 552 BP.
XX
AC AAX35375;
XX
DT 16-JUL-1999 (first entry)
XX
DE SFQ ID 50 of WO9916889.
XX
KW Angiostatin; endostatin; interferon; thrombospondin;
KW interferon-inducible protein; platelet factor 4; anti-angiogenic;
KW anti-tumor; multifunctional protein; angiogenic-mediated disease;
KW cancer; diabetic retinopathy; macular degeneration; arthritis;
XX tumor cell production; ss.
OS Homo sapiens.
XX

PN WO9916889-A1.
XX
PD 08-APR-1999.
XX
PF 30-SEP-1998; 98WO-US20464.
XX
PR 01-OCT-1997; 97US-0060609.
XX
PA (SEAR) SEARLE & CO G D.
XX
PI Bolanowski MA, Caparon MH, Casperson GF, Gregory SA;
PI Klein BK, McKearn JP;
XX WPI; 1999-255098/21.
XX
PT New multifunctional proteins useful for treating angiogenic-mediated
PT diseases
XX
PS Disclosure; Page 85; 121pp; English.
XX
CC The specification describes multifunctional proteins which comprise
CC combinations of angiostatin, endostatin, interferon, thrombospondin,
CC interferon-inducible protein and platelet factor 4, and have
CC anti-angiogenic and/or anti-tumor activity. The multifunctional protein
CC may exhibit useful properties such as having similar or greater
CC biological activity when compared to a single factor or by having
CC improved half-life or decreased adverse side effects, or a combination
CC of these properties. The proteins can be used for treating an
CC angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular
CC degeneration, or arthritis. They can also be used for inhibiting the
CC production of tumor cells (characteristic of lung, breast, ovarian,
CC prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma,
CC hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor
CC growth. The present sequence is used in the course of the invention.
XX
SQ Sequence 552 BP; 83 A; 196 C; 179 G; 94 T; 0 other;

Query Match 4.7%; Score 26; DB 20; Length 552;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGTTCAGCAGCGCG 113
Db 88 GACTTCCAGTGTTCAGCAGCGCG 113

RESULT 15
AAC62023
ID AAC62023 standard; DNA; 552 BP.
XX
AC AAC62023;
XX
DT 06-MAR-2001 (first entry)
XX
DE Nucleotide sequence of human endostatin encoded by plasmid pMALCH15.
XX
KW Streptomyces sp. strain C5; SnpA; S. venezuelae; alpha-amylase;
KW endostatin; cancer; tumour growth; angiogenesis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..552
FT /tag= a
FT /product= "endostatin"
XX
XX WO200060945-A1.
XX
PD 19-OCT-2000.
XX
PF 12-APR-2000; 2000WO-US09747.
XX
PR 13-APR-1999; 99US-0129084.


```

; FILING DATE: 22-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07880/033001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-956-307B-10

Query Match 3.2%; Score 18; DB 3; Length 288;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCGGCACCTTCGGGCC 144
Db 54 GCGGCACCTTCGGGCC 71

RESULT 7
US-08-956-307B-9/c
; Sequence 9, Application US/08956307B
; Patent No. 6090911
; GENERAL INFORMATION:
; APPLICANT: Petka, Wendy A.
; APPLICANT: Tirrell, David A.
; APPLICANT: Kevin P. McGrath
; TITLE OF INVENTION: REVERSIBLE HYDROGELS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,307B
; FILING DATE: 22-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07880/033001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 292 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA

US-08-956-307B-9

Query Match 3.2%; Score 18; DB 3; Length 292;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCGGCACCTTCGGGCC 144
Db 235 GCGGCACCTTCGGGCC 218

RESULT 8
US-09-449-293-3
; Sequence 3, Application US/09449293
; Patent No. 6267954
; GENERAL INFORMATION:
; APPLICANT: Abitbol, Marc
; APPLICANT: Uteza, Yves
; APPLICANT: Menasche, Maurice
; APPLICANT: Bossard, Carine
; APPLICANT: Van Den Berghe, Loic
; APPLICANT: Bonnel, Sebastien
; APPLICANT: Prats, Herve
; APPLICANT: Honiger, Jiri
; APPLICANT: Neuner-Jehle, Martin
; TITLE OF INVENTION: INTRAOCULAR TRANSPLANTATION OF ENCAPSULATED CELLS
; FILE REFERENCE: 8076.202US01
; CURRENT APPLICATION NUMBER: US/09/449,293
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Rattus rattus
US-09-449-293-3

Query Match 3.2%; Score 18; DB 3; Length 558;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 GACGCAGAGATGTCCTG 327
Db 313 GACGCAGAGATGTCCTG 330

RESULT 9
US-09-775-325-3
; Sequence 3, Application US/09775325
; Patent No. 6500449
; GENERAL INFORMATION:
; APPLICANT: Abitbol, Marc
; APPLICANT: Uteza, Yves
; APPLICANT: Menasche, Maurice
; APPLICANT: Bossard, Carine
; APPLICANT: Van Den Berghe, Loic
; APPLICANT: Bonnel, Sebastien
; APPLICANT: Prats, Herve
; APPLICANT: Honiger, Jiri
; APPLICANT: Neuner-Jehle, Martin
; TITLE OF INVENTION: INTRAOCULAR TRANSPLANTATION OF ENCAPSULATED CELLS
; FILE REFERENCE: 8076.202US01
; CURRENT APPLICATION NUMBER: US/09/775,325
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: US 09/449,293
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Rattus rattus
US-09-775-325-3
```

```
Query Match      3.2%; Score 18; DB 4; Length 558;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 GACGGCAGAGATGTCCTG 327
    |||||
Db 313 GACGGCAGAGATGTCCTG 330

RESULT 10
US-08-985-526-37
; Sequence 37, Application US/08985526
; Patent No. 6080728
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA
; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Connolly, Bove, Lodge, & Hutz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/985,526
; APPLICATION NUMBER: US/08/985,526
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorrow Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 565 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-985-526-37

Query Match      3.2%; Score 18; DB 3; Length 565;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 GACGGCAGAGATGTCCTG 327
    |||||
Db 320 GACGGCAGAGATGTCCTG 337

RESULT 11
US-09-561-500-12
; Sequence 12, Application US/09561500
; Patent No. 6342219
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002500
; CURRENT APPLICATION NUMBER: US/09/561,500
; PRIOR FILING DATE: 2000-04-28
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
```

```
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: OLIGONUCLEOTIDE
; NAME/KEY: CDS
; LOCATION: (1)..(573)
; US-09-561-500-12

Query Match      3.2%; Score 18; DB 4; Length 573;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 GACGGCAGAGATGTCCTG 327
    |||||
Db 331 GACGGCAGAGATGTCCTG 348

RESULT 12
US-09-561-108-12
; Sequence 12, Application US/09561108
; Patent No. 6342221
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002584
; CURRENT APPLICATION NUMBER: US/09/561,108
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: OLIGONUCLEOTIDE
; NAME/KEY: CDS
; LOCATION: (1)..(573)
; US-09-561-108-12

Query Match      3.2%; Score 18; DB 4; Length 573;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 GACGGCAGAGATGTCCTG 327
    |||||
Db 331 GACGGCAGAGATGTCCTG 348

RESULT 13
US-09-561-526-12
; Sequence 12, Application US/09561526
; Patent No. 6416758
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE KITS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002586
; CURRENT APPLICATION NUMBER: US/09/561,526
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
```

SEQ ID NO 12
LENGTH: 573
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
NAME/KEY: CDS
LOCATION: (1)..(573)
US-09-561-526-12

Query Match 3.2%; Score 18; DB 4; Length 573;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 GACGGCAGAGATGTCCTG 327
DB 331 GACGGCAGAGATGTCCTG 348

RESULT 14
US-09-561-499-12
Sequence 12, Application US/09561499
Patent No. 6524583
GENERAL INFORMATION:
APPLICANT: Philip E. Thorpe
APPLICANT: Rolf A. Brekken
TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF
FILE REFERENCE: 4001.002582
CURRENT APPLICATION NUMBER: US/09/561,499
CURRENT FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/131,432
PRIOR FILING DATE: 1999-04-28
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 573
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
NAME/KEY: CDS
LOCATION: (1)..(573)
US-09-561-499-12

Query Match 3.2%; Score 18; DB 4; Length 573;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 GACGGCAGAGATGTCCTG 327
DB 331 GACGGCAGAGATGTCCTG 348

RESULT 15
US-08-956-307B-12
Sequence 12, Application US/08956307B
Patent No. 6090911
GENERAL INFORMATION:
APPLICANT: Petka, Wendy A.
APPLICANT: Tirrell, David A.
APPLICANT: Kevin P. McGrath
TITLE OF INVENTION: REVERSIBLE HYDROGELS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,307B
FILING DATE: 22-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07880/033001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 774 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-956-307B-12

Query Match 3.2%; Score 18; DB 3; Length 774;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GCCGGCACCTTCGGGGCC 144
DB 54 GCCGGCACCTTCGGGGCC 71

Search completed: August 17, 2003, 19:49:28
Job time : 65.5426 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 16:23:08 ; Search time 1617.68 seconds
(without alignments)
8338.472 Million cell updates/sec

Title: US-09-938-391-3
Perfect score: 555
Sequence: 1 cacaccaccaggacttcca.....tgacctcttcccaagtag 555

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estcom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	8.5	279	9	AW464343
2	47	8.5	387	10	BF601253
3	47	8.5	551	10	BF074459
4	41	7.4	657	14	CB444165

C	5	31	5.6	409	14	CB537919	CB537919 775852 MA
C	6	31	5.6	623	14	CB424313	CB424313 598557 MA
C	7	29	5.2	356	10	BF708656	BF708656 MI-P-AVO-
	8	29	5.2	451	10	EG383960	EG383960 302432 MA
	9	29	5.2	491	10	EG383970	EG383970 302444 MA
	10	26	4.7	390	9	AV660284	AV660284 AV660284
	11	26	4.7	408	10	BF602228	BF602228 267408 MA
	12	26	4.7	483	12	EG823096	EG823096 602726433
	13	26	4.7	485	9	AV727686	AV727686 AV727686
	14	26	4.7	505	9	AV727879	AV727879 AV727879
	15	26	4.7	556	14	CB216999	CB216999 NISC nq11
	16	26	4.7	612	12	BM768821	BM768821 K-EST0051
	17	26	4.7	641	9	AUI27824	AUI27824 AUI27824
	18	26	4.7	683	12	BM683067	BM683067 UI-E-EOI-
	19	26	4.7	703	13	BU615520	BU615520 UI-H-FGO-
	20	26	4.7	707	10	BE908201	BE908201 601500458
	21	26	4.7	709	14	CB443805	CB443805 694911 MA
	22	26	4.7	715	9	AUI25614	AUI25614 AUI25614
	23	26	4.7	757	10	BE906253	BE906253 601502237
	24	26	4.7	832	10	EG387051	EG387051 602454749
	25	26	4.7	846	13	BU540812	BU540812 AGENCOURT
	26	26	4.7	866	13	EX451426	EX451426 BX451426
	27	26	4.7	881	14	CD105862	CD105862 AGENCOURT
	28	26	4.7	884	12	BI161007	BI161007 602865213
	29	26	4.7	929	13	BQ672290	BQ672290 AGENCOURT
	30	26	4.7	934	13	BU944034	BU944034 AGENCOURT
	31	26	4.7	944	13	BU859398	BU859398 AGENCOURT
	32	26	4.7	947	13	BU556872	BU556872 AGENCOURT
	33	26	4.7	979	13	BQ673186	BQ673186 AGENCOURT
	34	26	4.7	1093	13	BQ723254	BQ723254 AGENCOURT
	35	26	4.7	1121	12	BM564439	BM564439 AGENCOURT
	36	25	4.5	402	9	AV696283	AV696283 AV696283
	37	25	4.5	445	9	AI809446	AI809446 wf70907.x
	38	24	4.3	1131	13	BQ897787	BQ897787 AGENCOURT
	39	23	4.1	867	13	BQ227326	BQ227326 603798785
	40	22	4.0	323	10	BI187708	BI187708 BI187708
	41	22	4.0	336	9	AW141057	AW141057 EST291080
	42	22	4.0	424	14	CB796474	CB796474 AMGNNUC-S
	43	22	4.0	518	14	CB716008	CB716008 AMGNNUC-S
	44	22	4.0	570	12	BI288582	BI288582 UI-R-DKO-
	45	22	4.0	577	12	BM488074	BM488074 pgmzn.pk0

ALIGNMENTS

RESULT 1	AW464343	279 bp	mRNA	linear	EST 24-FEB-2000
LOCUS	BP230015B10C11	Soares normalized	bovine placenta	Bos taurus	CDNA
DEFINITION	Clone BP230015B10C11 5', mRNA sequence.				
ACCESSION	AW464343				
VERSION	AW464343.1	GI:7034511			
KEYWORDS	EST.				
SOURCE	Bos taurus (cow)				
ORGANISM	Bos taurus				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.				
REFERENCE	1 (bases 1 to 279)				
AUTHORS	Lewin,H.A., Soares,M.B., Rebeiz,M., Pardini,J., Liu,L. and Larson J.H.				
TITLE	Bovine ESTs				
JOURNAL	Unpublished				
COMMENT	Contact: Lewin, H. A. W. M. Keck Center for Comparative and Functional Genomics University of Illinois at Urbana-Champaign 340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL 61801, USA Tel: 217 333 5998 Fax: 217 244 5617 Email: h-lewin@uiuc.edu Funding for cattle EST sequencing was provided by the USDA National				

Research Initiative, Animal Genome Resource Grant AG 99-3205-8534 to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED from Washington University Genome Center. Vector Trimmi 9: Cross match from Washington University Genome Center PHRAP suite. Sequences submitted are vector free and at least 200 bp in length. PCR Primers
 FORWARD: TAATACGACTCACTATAGG
 BACKWARD: ATTAACCTCACTAAG
 Insert length: 279 Std Error: 0.00
 Plate: BP230019B10 row: C column: 11
 Seq primer: AGCGATACAAATTCACACAGGA
 High quality sequence stop: 279.

FEATURES

source
 1. .279
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /clone="BP230015B10C11"
 /sex="female"
 /lab_host="DH10B"
 /clone_lib="Soares normalized bovine placenta"
 /note="Organ: placenta; Vector: pT73Pac; Site 1: EcoRI; Site 2: NotI; The cDNA library was contributed by the Soares laboratory and it was constructed and normalized as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1986), Genome Research 6(9): 791-806."

BASE COUNT 35 a 106 c 86 g 52 t
 ORIGIN

Query Match 8.5%; Score 47; DB 9; Length 279;
 Best Local Similarity 100.0%; Pred. No. 3.4e-11;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 94 CAGTGTCTCCAGCAGCGCGCGCGGCTGGCGGACCTTCG 140
 Db 24 CAGTGTCTCCAGCAGCGCGCGCGGCTGGCGGACCTTCG 70

RESULT 2

BF601253
 LOCUS
 DEFINITION 387 bp mRNA linear EST 25-APR-2001
 266182 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION
 BF601253
 VERSION
 BF601253.1 GI:11698475
 EST.
 SOURCE
 Bos taurus (cow)
 ORGANISM
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 387)
 Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
 G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
 Perteau,G., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and
 Keele,J.W.

TITLE
 Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle

JOURNAL
 MEDLINE
 PUBMED
 21180013
 11282978

COMMENT
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -minmatch 12 options.
 PCR Primers
 FORWARD: AGGAAACAGCTATGACCAT
 BACKWARD: GTTTTCCAGTCACGACG

Plate: 40 row: H column: 7
 Seq primer: ATTAGGTGACACTATAG.
 Location/Qualifiers
 1. .387
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 3BOV"
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
 Library made from pooled tissue from marrow, alveolar
 macrophage, ovary, fetal semitendinosus muscle, and fetal
 longissimus muscle."

BASE COUNT 52 a 140 c 131 g 64 t
 ORIGIN

Query Match 8.5%; Score 47; DB 10; Length 387;
 Best Local Similarity 100.0%; Pred. No. 3.5e-11;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 94 CAGTGTCTCCAGCAGCGCGCGGCTGGCGGACCTTCG 140
 Db 7 CAGTGTCTCCAGCAGCGCGCGGCTGGCGGACCTTCG 53

RESULT 3
 BF074459
 LOCUS
 DEFINITION 551 bp mRNA linear EST 25-APR-2001
 221883 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION
 BF074459
 VERSION
 BF074459.1 GI:10867970
 EST.
 SOURCE
 Bos taurus (cow)
 ORGANISM
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 551)
 Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
 G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
 Perteau,G., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and
 Keele,J.W.

TITLE
 Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle

JOURNAL
 MEDLINE
 PUBMED
 21180013
 11282978

COMMENT
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -minmatch 12 options.
 PCR Primers
 FORWARD: AGGAAACAGCTATGACCAT
 BACKWARD: GTTTTCCAGTCACGACG

Plate: 81 row: F column: 9
 Seq primer: ATTAGGTGACACTATAG.
 Location/Qualifiers
 1. .551
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 2BOV"
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
 Library made from pooled tissue from testis, thymus,


```

FEATURES
  source
    Location/Qualifiers
      1. .623
        /organism="Bos taurus"
        /mol_type="mRNA"
        /db_xref="taxon:9913"
        /tissue_type="pooled"
        /lab_host="DH10B"
        /clone_lib="MARC 6BOV"
        /note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."
BASE COUNT 105 a 199 c 193 g 126 t
ORIGIN
  206 TGCCCGTCGTCACCTCAGGCGAGGTGCT 236
  Db 572 TGCCCGTCGTCACCTCAGGCGAGGTGCT 542

RESULT 7
LOCUS BF708656/c 356 bp mRNA linear EST 02-JAN-2001
DEFINITION MI-P-AYO-nq-g-09-0-UI.s1 MI-P-AYO Sus scrofa cDNA clone
ACCESSION BF708656
VERSION BF708656.1 GI:12008133
KEYWORDS EST.
SOURCE Sus scrofa (pig)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
  REFERENCE
    1 (bases 1 to 356)
    Ronald M.F. Lennon G. and Soares M.B.
  AUTHORS
    Normalization and subtraction: two approaches to facilitate gene
    discovery
  JOURNAL
    Genome Res. 6 (9), 791-806 (1996)
  MEDLINE
    97044477
  PUBMED
    889548
  COMMENT
    Contact: Tuggle CK
    Molecular Genetics Laboratory, Department of Animal Science
    Iowa State University
    201 Kildee Hall, Ames, IA 50011-3150, USA
    Tel: 5152944252
    Fax: 5152942401
    Email: cktuggle@iastate.edu
    The sequence contained an oligo-dT track that was present in the
    oligonucleotide that was used to prime the synthesis of first
    strand cDNA and therefore this may represent a bonafide poly A
    tail. The sequence tag present in the cDNA between the NotI site
    and the oligo-dT track served to verify it as a clone from the
    non-normalized placenta library cDNA library preparation: M.B.
    Soares Lab, University of Iowa EST sequencing: M.B. Soares Lab,
    University of Iowa Clone distribution: clones will be available
    through Research Genetics (www.resgen.com)
    Seq primer: M13 Forward
    POLYA=Yes.

FEATURES
  source
    Location/Qualifiers
      1. .356
        /organism="Sus scrofa"
        /mol_type="mRNA"
        /db_xref="taxon:9823"
        /clone="MI-P-AYO-nq-g-09-0-UI"
        /lab_host="DH10B (Life Technologies)"
        /clone_lib="MI-P-AYO"
        /note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: EcoRI; The MI-P-AYO
library is derived from placenta. For a detailed
description of the library from which this clone was
derived, please visit our web site at
http://pigest.genome.iastate.edu/. The procedure used to
create this library has been previously described (Ronald
M. Lennon and Soares, Genome Research 6:791-806, 1996)
TAG LIB=MI-P-AYO
TAG TISSUE=placenta
TAG_SEQ=ATTGG"
BASE COUNT 61 a 109 c 101 g 84 t 1 others
ORIGIN
  292 GCCCGCATCTCTCTTCGACGGCAGAGA 320
  Db 340 GCCCGCATCTCTCTTCGACGGCAGAGA 312

Query Match 5.2%; Score 29; DB 10; Length 356;
Best Local Similarity 100.0%; Pred. No. 0.0082;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 292 GCCCGCATCTCTCTTCGACGGCAGAGA 320
  ||||||||||||||||||||||||||||
Db 340 GCCCGCATCTCTCTTCGACGGCAGAGA 312
  ||||||||||||||||||||||||||||

RESULT 8
LOCUS BG383960 451 bp mRNA linear EST 12-MAR-2001
DEFINITION 302432 MARC IPiG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BG383960
VERSION BG383960.1 GI:13308432
KEYWORDS EST.
SOURCE Sus scrofa (pig)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
  REFERENCE
    1 (bases 1 to 451)
    Fahrenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,
    Vallet, J., Wise, T., Rohrer, G.A., Perlea, G., Sultana, R., Quackenbush
    J. and Keeler, J.W.
  AUTHORS
    Porcine Gene discovery by normalized cDNA-library sequencing and
    EST cluster assembly
  JOURNAL
    Mamm. Genome 13 (8), 475-478 (2002)
  MEDLINE
    22213789
  PUBMED
    12226715
  COMMENT
    Contact: Smith TPL
    USDA, ARS, US Meat Animal Research Center
    PO Box 166, Clay Center, NE 68933-0166, USA
    Tel: 402 762 4366
    Fax: 402 762 4390
    Email: smith@email.marc.usda.gov
    Single pass sequencing. Bases called and alt trimmed with phred
    v0.980904.e. Vector identified by cross_match with the -minscore 18
    and -minmatch 12 options.
    PCR Primers
    FORWARD: AGGAACACAGCTATGACCAT
    BACKWARD: GTTTTCCAGTCACGACG
    Plate: 89 row: I column: 13
    Seq primer: ATTAGTGACACTATAG.
    Location/Qualifiers
      1. .451
        /organism="Sus scrofa"
        /mol_type="mRNA"
        /db_xref="taxon:9823"
        /tissue_type="pooled"
        /lab_host="DH10B"
        /clone_lib="MARC IPiG"
        /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
BASE COUNT 75 a 153 c 152 g 71 t
ORIGIN
  292 GCCCGCATCTCTCTTCGACGGCAGAGA 320
  ||||||||||||||||||||||||||||
QY 292 GCCCGCATCTCTCTTCGACGGCAGAGA 320
  ||||||||||||||||||||||||||||

Query Match 5.2%; Score 29; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 0.0084;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```
REFERENCE
AUTHORS    1 (bases 1 to 505)
            Gu, Y., Peng, Y., Song, H., Huang, Q., Yang, Y., Gao, G., Xiao, H., Xu, X.,
            Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu
            S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R.,
            Chen, J., Chen, Z. and Han, Z.
TITLE      Homo sapiens cDNA HTC clones
JOURNAL    Unpublished
COMMENT    Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.

FEATURES
source     1..505
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="HTCARG03"
            /tissue_type="Hypothalamus"
            /dev_stage="Adult"
            /lab_host="SOLR"
            /clone_lib="HTC"
            /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
            XhoI"
BASE COUNT 77 a 212 c 152 g 61 t      3 others
ORIGIN
Query Match      4.7%; Score 26; DB 9; Length 505;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGCTTCCAGCAGCGCG 113
      |||
Db 394 GACTTCCAGTGCTTCCAGCAGCGCG 419

RESULT 15
CB216999
LOCUS      NISC.nq11c06.y1 NICHD_HS_Ut2 Homo sapiens cDNA clone IMAGE:5938570
DEFINITION 5', mRNA sequence.
ACCESSION  CB216999
VERSION    CB216999.1 GI:28265191
KEYWORDS  EST.
SOURCE     Homo sapiens (human)
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 556)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL    Unpublished
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            CDNA Library Preparation:
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC)
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            info@image.llnl.gov
            Plate: LLAM13167 row: F column: 11
            Seq primer: M13RPI reverse primer (ABI).
            Location/Qualifiers
            1..556
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5938570"

FEATURES
source     1..556
            /sex="female"
            /tissue_type="normal endometrium, mid-secretory phase,
            cycle day 23"
            /lab_host="DH10B (Tl-resistant)"
            /clone_lib="NICHD_HS_Ut2"
            /note="Organ: uterus; Vector: pCMV-SPORT6.1.cdb (ResGen,
            Invitrogen Corporation); Site_1: NotI; Site_2: EcoRV;
            Cloned unidirectionally from microquantity amounts of mRNA
            from normal endometrial tissue (mid-secretory phase, cycle
            day 23). Average insert size 1.6 kb. Library constructed
            by ResGen (Invitrogen Corporation)."
BASE COUNT 78 a 236 c 163 g 79 t
ORIGIN
Query Match      4.7%; Score 26; DB 14; Length 556;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GACTTCCAGTGCTTCCAGCAGCGCG 113
      |||
Db 299 GACTTCCAGTGCTTCCAGCAGCGCG 324

Search completed: August 17, 2003, 19:46:44
Job time : 1619.68 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 13:17:18 ; Search time 2786.63 Seconds
(without alignments)
8147.780 Million cell updates/sec

Title: US-09-938-391-3

Perfect score: 555

Sequence: 1 cacaccaccaggacttcca.....tgactctcttctccaagtag 555

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fuh.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pin.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	555	100.0	555	6	AX399631	Sequence
2	555	100.0	829	6	AX399629	Sequence
3	419.6	75.6	2637	9	BC033715	Homo sapi
4	418	75.3	546	6	AR193165	Sequence
5	418	75.3	549	6	AX100086	Sequence
6	418	75.3	552	6	AX395662	Sequence
7	418	75.3	552	6	AX473835	Sequence
8	418	75.3	552	6	BD081407	Fused pro
9	418	75.3	786	6	AF282883	Homo sapi
10	418	75.3	900	6	AX370851	Sequence
11	418	75.3	2364	9	AK098216	Homo sapi
12	418	75.3	3394	6	AX409531	Sequence
13	418	75.3	3394	6	I51045	Sequence 4
14	418	75.3	3394	9	HUMCOL18AX	Human colla
15	418	75.3	5408	9	AF018082	Homo sapi
16	418	75.3	5529	9	AF018081	Homo sapi
17	416.4	75.0	564	9	AF416592	Homo sapi
18	411.6	74.2	555	9	AF184060	Homo sapi
19	410.8	74.0	534	6	AR193166	Sequence
20	410.8	74.0	537	6	AX100092	Sequence
21	398.4	71.8	650	6	AX370853	Sequence
22	379	68.3	851	10	AF189709	Rattus no
23	375.8	67.7	555	6	BD102800	Process f
24	375.8	67.7	555	6	BD128313	Endostati
25	375.8	67.7	558	6	AR268609	Sequence
26	375.8	67.7	565	6	AR100642	Sequence
27	375.8	67.7	565	6	E34073	Carrier/DNA
28	375.8	67.7	624	6	AX084406	Sequence
29	375.8	67.7	1615	10	MMU03714	Mus musculu
30	375.8	67.7	4437	10	MUSCOLLAG	Mus musculu
31	375.8	67.7	4802	10	BC043697	Mus muscu
32	374.2	67.4	624	6	AX128409	Sequence
33	374.2	67.4	4031	6	I51044	Sequence 1
34	374.2	67.4	4031	10	MUSCOLPRO	Mouse alpha
35	373.4	67.3	552	10	AF257775	Mus muscu
36	373.4	67.3	573	6	AR183471	Sequence
37	373.4	67.3	573	6	AR183520	Sequence
38	373.4	67.3	573	6	AR217279	Sequence
39	373.4	67.3	573	6	AR282751	Sequence
40	373.4	67.3	573	6	AX042272	Sequence
41	373.4	67.3	573	6	AX490666	Sequence
42	350.8	63.2	515	10	RNO236873	Rattus no
43	345	62.2	5279	5	AF083440	Gallus ga
44	251	45.2	5598	5	AB047066	Xenopus l
45	245.2	44.2	2037	5	DRE494837	Danio rer

ALIGNMENTS

RESULT 1	AX399631	Sequence 3 from Patent EP1191036.	555 bp	DNA	linear	PAT 06-JUN-2002
LOCUS	AX399631					
DEFINITION	Sequence 3 from Patent EP1191036.					
ACCESSION	AX399631					
VERSION	AX399631.1	GI:21335410				
KEYWORDS	Canis familiaris (dog)					
SOURCE	Canis familiaris					
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.					
REFERENCE	1	Sheppard, M.G. and Tong, X.				
AUTHORS		Methods and compositions for diagnosing and treating disorders				
TITLE		involving angiogenesis				

Qy	301	TTCTCTTTGACGGGAGAGATGTCTCTGACGACACCCCGCTGGCCCCGGAAGAGCGTGTGG	360
Db	301	TTCTCTTTGACGGGAGAGCGTCTCTGAGGACACCCACCTGGCCCCAGAGAGCGTGTGG	360
Qy	361	CACGGCTCCGACCCACAGCGGGCGCGCTGTGACCGACACTACTGCGAGACGTGGCGGACG	420
Db	361	CATGGCTCGAACCCACAGGGCGGAGGCTGACCGAGAGCTTACTGTGACACTGGCGGACG	420
Qy	421	GAGCCCCGGCGGGCCACCGGCGAGCGGTCTGTCTGCTGGCGGGCAGGCTCTCTGAGCAG	480
Db	421	GAGGCTCCCTGGGCGCACGGGGCCAGGCGCTCTCTCGTGTGGGGGGCAGGCTCTCTGGGGCAG	480
Qy	481	GAGCGCGGAGCTGCCCGCAGCCTTCTGTGTGTCTCTGCATCGAGAACAGCGTTCATGACC	540
Db	481	AGTCCCGGAGCTGCCATCACGCCTTACATCGTGTCTCTGCATTGAGAACAGCTTTCATGACT	540
Qy	541	TCCTTC	546
Db	541	GCCTCC	546

RESULT 5					
AX100086					
LOCUS	AX100086	549 bp	DNA	linear	PAT 02-APR-2001
DEFINITION	Sequence 4 from Patent WO0119989.				
ACCESSION	AX100086				
VERSION	AX100086.1	GI:13539061			
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1				
AUTHORS	Liang, H., Sim, K.L., Chang-Murad, A., Zhou, X., Madsen, J., Boerner, R.J., Bermejo, L.L., Mistry, F.R., Shepard, S.R. and Schrimsher, J.				
TITLE	Method of producing and purifying endostatin ² tm protein				
JOURNAL	Patent: WO 0119989-A 4 22-MAR-2001;				
	EntreMed, Inc. (US)				
FEATURES	Location/Qualifiers				
source	1. .549				

```

BASE COUNT
ORIGIN
82 a 196 c 178 g 93 t
/OW_XREF="taxon:9606"

```

Query Match	75.3%;	Score 418;	DB 6;	Length 549;
Best Local Similarity	85.3%;	Pred. No. 1.9e-46;		
Matches 466;	Conservative 0;	Mismatches 80;	Indels 0;	Gaps 0;

Qy	1	CACACCACACGAGGACTTCCAGCTGGTGTCTGCACTGTGTGGCCCTGGAACAGGCCGCGAGCGG	60
Db	1	CACAGCCACCGCGACTTCCAGCCGGTGCTCCACCTGGTTGGCTCTCAACAGGCCCTCTGTCA	60
Qy	61	GGCGGCATGCGAGGCATCCGGGAGCGGACTTCCAGTCTTCCAGCAGGCGCGCGCGCG	120
Db	61	GGCGGCATGCGGGGCATCCGGGGCGGACTTCCAGTCTTCCAGCAGGCGCGGGCGCGT	120
Qy	121	GGGCTGGCGGCACCTTCCGGGCGCTTCTCTGCTCGTGGGCTGCAGGACCTCTACAGCATC	180
Db	121	GGGCTGGGGGCACTTCCGGCGCTTCTCTGCTCGGCGCTTGCAGGACCTGTACAGCATC	180
Qy	181	GTGCGCGCGCGCGACCGCACCGGGGTGCCGCTCGTCAACCTCAGGACGAGGTGCTTTC	240
Db	181	GTGCGCGGTGCCGACCGCGGACGCGTGCACCTGTCAACCTCAGGACGAGTCTGTTTT	240
Qy	241	CCCACTGGGAGGCTTATTCTCGGGCTCCGAGGCCAGCTGAAGCCCGGGGCCCGCATC	300
Db	241	CCCACTGGGAGGCTCTGTTCTCAGGCTCTGAGGTCGCGCTGAAGCCCGGGGCGCAGCATC	300
Qy	301	TTCTCTTTCAGCGGAGAGATGTCTCTGACGACCCCGCGCTGGCCCCCGAAGAGCGTGTGG	360

301	TTTCTCTTTGACGGCAAGGAGCTCTGAGGCA	CCCCACCTGGCCCCAGAGAGCGTGTGG	360
Db			
361	CACGGCTTCGAGACCCCAAGCGGGCGCGCTT	GACCGACAGCTACTCGGAGACGTGTGCGGACG	420
Qy			
361	CATGGCTCGGACCCCAACGGGCGCAGGCT	TGACCGAGAGCTACTGTGAGACGTGTGCGGACG	420
Db			
421	GAGCCCCGGCGGGCCACCGGGCAGGCGT	CGTCTGCTGTGGGGGCGAGGCTGCTGGAGCAG	480
Qy			
421	GAGGCTCCCTCGGCACCGGCCAGGGCCT	CTCTCGTGTGGGGGCGAGGCTCTCTGGGCGAG	480
Db			
481	GAGCGCGGAGCTGCGGCCACGCCTTCGT	GGTGTCTGTGATTCGAGAAACAGCGTTCATGACC	540
Qy			
481	AGTGCCCGGAGCTGCGCATACGCGCT	TACATCGTGTCTGTGCATTGAGAACAGCTTCATGACT	540
Db			
541	TCCTTC	546	
Qy			
541	GCCTCC	546	
Db			

RESULT 6					
AX395662	AX395662	Sequence 30 from Patent WO0193897.	552 bp	DNA	linear
LOCUS					
DEFINITION					
AX395662					
ACCESSION					
AX395662					
VERSION					
AX395662.1					
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				

REFERENCE
1. Sim, K. L. and Macdonald, N. J.
AUTHORS
TITLE Angiotensin and endostatin binding proteins and methods of use
JOURNAL Patent: WO 0193897-A 30 13-DEC-2001;
Entrevued, Inc. (US)
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

FEATURES
source
Location/Qualifiers
1..552
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT
83 a 196 c 179 g 94 t
ORIGIN

```

Query Match	75.3%	Score 418;	DB 6;	Length 552;
Best Local Similarity	85.3%;	Pred. No. 1.9e-46;		
Matches 466;	Conservative	0;	Mismatches 80;	Indels 80;
				Gaps 0;

Qy	1	CACACCCACAGGACTTTCAGCTGTGTCTCGACCTGTGTGGCCCTGAAACAGCCCGCAGCGC	60
Db	1		
Db	1	CACAGCCACCGCGACTTTCAGCGCGGTCTCCACCTGTGTGCGCTCAACAGCCCCCTGTCA	60
Qy	61	GGCGGCATGCGAGGCATCCGGGAGCGGACTTCCAGTGCTTCCAGCAGGCGCGCGCGCG	120
Db	61		
Qy	121	GGGCTGCGCGGCACTTTCGCGGCTTCTCTGCTCGCGGTGCGAGGACCTCTACAGCATC	180
Db	121	GGGCTGCGCGGCACCTTTCGCGGCTTCTCTGCTCGCGCTGCGAGGACCTGTACAGCATC	180
Qy	181	GTGCGCGCGCCACACGCGACCGGGTGCCCGTGTCAACCTCAGGGACGAGGTGCTCTTC	240
Db	181		
Qy	241	CCAGCTGGGAGGCTTATTCTCGGGCTCCGAGGGCCAGCTGAAGCCCGGGGCCCGCATC	300
Db	241		
Qy	301	TTCTCTTTTGACGCGCAGAGATGTCTGCGACACCCCGCTGGCCCCCGGAAGAGCGTGTGG	360
Db	301	TTCTCTTTTGACGCGCAGGACGTCTTGAGGCACCCCACTGGCCCCCAGAGAGCGTGTGG	360
Qy	361	CACGGCTCCGACCCCGAGCGGGCGCGCTGACCGACAGCTACTGCGAGACGTGGCGGACG	420
Db			

```

Db      361  CATGGCTCGGACCCCAACCGGGCGCAGGCTGACCGAGAGTACTGTGTAGACGTGTGGGACG 420
QY      421  GAGGCCCCCGCGCCACCGGGCAGGCGTCTCGCTCGCTCGCGCGCAGGCTGCTGTAGACG 480
Db      421  GAGGCTCCCTCGGCGACGCGCCAGGCTCTCGCTGTGGGGGCGAGCTCTGTGGGCGAG 480
QY      481  GAGGCGCGAGCTCGCGCCACGCGCTCTGTGTGTGTCTGTGCATCGAGAACAGGCTCATGACC 540
Db      481  AGTGGCGGAGCTGCCATCACGCTACATCGTGTCTGTGATTTGAGAACAGCTTCATGACT 540
QY      541  TCCTTC 546
Db      541  GCCTCC 546

RESULT 7
LOCUS      AX473835
DEFINITION Sequence 2 from Patent WO0230982.
ACCESSION AX473835
VERSION   AX473835.1 GI:22208005
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE
AUTHORS   Sim,K.L. and Liang,H.
TITLE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL   Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
FEATURES
source    1..552
           Location/Qualifiers
           1..552
           /organism="Homo sapiens"
           /mol_type="genomic DNA"
           /db_xref="taxon:9606"
BASE COUNT 83 a 196 c 179 g 94 t
ORIGIN

Query Match 75.3%; Score 418; DB 6; Length 552;
Best Local Similarity 85.3%; Pred. No. 1.9e-46;
Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY      1  CACACCCACGAGACTTCCAGCTGTGTGTCACCTGTGTGCGCCCTGAACAGCCGCGAGCG 60
Db      1  CACAGCCACCGGACTTCCAGCGGTGTCTCCACCTGTGTGCGCTCAACAGCCCTGTCTCA 60
QY      61  GCGCGCATCGAGGATCCGGGAGCGGACTTCCAGTGTCTTCCAGAGCGCGCGCGCG 120
Db      61  GCGCGCATCGGGGATCCGGGGCGGACTTCCAGTGTCTTCCAGTGTCTTCCAGAGCGCGCGTG 120
QY      121  GGGCTGGCGGACCTTCCGGGCGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Db      121  GGGCTGGCGGACCTTCCGGCGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY      181  GTGCGCGCGCGCGACCGAGCGGGGTGCGCGTCCGTCGTCACCTCAGGACGAGTGTCTTTC 240
Db      181  GTGCGCGGTGCGCGACCGCGCGCGCGTCCGTCGTCACCTCAGGACGAGTGTGTGT 240
QY      241  CCCAGTGGAGGCGCTTATTTCTGGGCTCCGAGGCGCGAGCTGAAGCCCGGGGCCCGCATC 300
Db      241  CCCAGTGGAGGCTGTGTCTCAGGCTCTGAGGCTCGCTGAAGCCCGGGGCCCGCATC 300
QY      301  TTCTCTTTTCGAGGAGATGCTCTGAGACCGCGCGCGCGCTGCGCGCGCGAGAGCGTGTGG 360
Db      301  TTCTCTTTTCGAGGAGATGCTCTGAGGACCGCGCGCGCGCTGCGCGCGCGAGAGCGTGTGG 360
QY      361  CACGGCTCCGACCCAGCGGGCGCGCTGACCGACAGCTACTGCGAGACGTGGCGGACG 420
Db      361  CATGCTCGGACCCCAACCGGGCGGAGGCTGACCGAGAGTACTGTGTAGACGTGGCGGACG 420
QY      421  GAGGCCCCCGCGCCACCGGGCAGGCGTCTCGCTCGCTCGCGCGCAGGCTGCTGTAGACG 480

```

```

Db      421  GAGGCTCCCTCGGCGCACCGGGCGCAGGCTCTCGCTGTGTGGGGCGCAGGCTCTGTGGGCGAG 480
QY      481  GAGGCGCGAGCTCGCGCCACCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Db      481  AGTGGCGGAGCTGCCATCACGCTACATCGTGTCTGTGATTTGAGAACAGCTTCATGACT 540
QY      541  TCCTTC 546
Db      541  GCCTCC 546

RESULT 8
LOCUS      BD081407
DEFINITION Fused protein containing angiotensin component and utilization thereof in antitumor therapy.
ACCESSION BD081407
VERSION   BD081407.1 GI:22627010
KEYWORDS  JP 2001518304-A/50.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE
AUTHORS   Bolanowski,M.A., Caparon,M.H., Casperson,G.F., Gregory,S.A., Klein,B.K. and McKeown,J.P.
TITLE     Fused protein containing angiotensin component and utilization thereof in antitumor therapy
JOURNAL   Patent: JP 2001518304-A 50 16-OCT-2001;
COMMENT   GD SEARLE AND CO
           OS Homo sapiens (human)
           PN JP 2001518304-A/50
           PD 16-OCT-2001
           PP 30-SEP-1998 JP 20005113958
           PR 01-OCT-1997 US 60/060609
           PI MARK A BOLANOWSKI,MAIRE H CAPARON,GERALD F CASPERSON,SUSAN A GREGORY, PI BARBARA K KLEIN,JOHN P MCKEARN
           PC C12N15/09,A61K38/00,A61K48/00,A61P9/10,A61P35/00,C07K14/52, PC C07K14/56
           CC Fused protein containing angiotensin component and utilization thereof in antitumor therapy
           FH Key source
           FT source
           FT Location/Qualifiers
           1..552
           /organism="Homo sapiens"
           /mol_type="genomic DNA"
           /db_xref="taxon:9606"
BASE COUNT 83 a 196 c 179 g 94 t
ORIGIN

Query Match 75.3%; Score 418; DB 6; Length 552;
Best Local Similarity 85.3%; Pred. No. 1.9e-46;
Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY      1  CACACCCACGAGACTTCCAGCTGTGTGTCACCTGTGTGCGCCCTGAACAGCCGCGAGCG 60
Db      1  CACAGCCACCGGACTTCCAGCGGTGTCTCCACCTGTGTGCGCTCAACAGCCCTGTCTCA 60
QY      61  GCGCGCATCGAGGATCCGGGAGCGGACTTCCAGTGTCTTCCAGAGCGCGCGCGCG 120
Db      61  GCGCGCATCGGGGATCCGGGGCGGACTTCCAGTGTCTTCCAGTGTCTTCCAGAGCGCGCGTG 120
QY      121  GGGCTGGCGGACCTTCCGGGCGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Db      121  GGGCTGGCGGACCTTCCGGCGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY      181  GTGCGCGCGCGCGACCGAGCGGGGTGCGCGTCCGTCGTCACCTCAGGACGAGTGTCTTTC 240

```


Db 457 GGGCTGGCGGGAACCTTCCGGCGCTTCTCTGCTCGCGCTGCAGACCTGTACAGCATC 516
Qy 181 GTGCGCGCGCGACCGGCGCGCGCTGCTCAACTCAGGACGAGGTGCTCTTC 240
Db 517 GTGCGCGGTGGCGACCGCGCGCGCTGCTCAACTCAGGACGAGGTGCTCTTC 576
Qy 241 CCCAGCTGGGAGGCTTATTTCTCGGGCTCCGAGGCGCGCTGAAGCCCGGGCGCGCATC 300
Db 577 CCAGCTGGGAGGCTTCTGTTCTCAGGCTCTGAGGCTCCGCTGAAGCCCGGGCGCGCATC 636
Qy 301 TTCTCTTTTCAGCGGAGAGATGCTTCGAGCACCCTGCGCGCGCGCGGAGAGCGGTGG 360
Db 637 TTCTCTTTTCAGCGGAGAGCGTCTGAGGACACCCACCTGCGCGCGCGGAGAGCGGTGG 696
Qy 361 CACGGCTCGACCCCGAGCGCGCGCTGACCGACAGCTACTGCGAGACGTGGCGGACG 420
Db 697 CATGGCTCGACCCCGAGCGCGCGCTGACCGAGAGCTACTGCGAGCGGTGGCGGACG 756
Qy 421 GAGGCGCGCGCGCGCGCGCGCTGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db 757 GAGGCTCTCTCGCGCGCGCGCGCGCGCTGCTGCTGCTGCGCGCGCGCGCGCGCGCG 816
Qy 481 GAGGCGCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 817 AGTGGCGCGAGGCTGCGCATCAGCGCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 876
Qy 541 TCCTTC 546
Db 877 GCCTCC 882

RESULT 11

AK098216 2364 bp mRNA linear PRI 15-JUL-2002
LOCUS Homo sapiens cDNA FLJ40897 fis, clone UTERU2002964, highly similar
DEFINITION to Homo sapiens type XVIII collagen (COL18A1) mRNA.

AK098216
VERSION AK098216.1 GI:21758185

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Oshima,A., Takahashi-Fujii,A., Tanase,T., Inose,N., Takeuchi,K.,
Arita,M., Musashino,K., Yuuki,H., Haru,H., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Sugiyama,A., Kawakami,B.,
Suzuki,Y., Sugano,S., Negahori,K., Masuho,Y., Nagai,K. and
Isogai,T.
NEDO human cDNA sequencing project

Unpublished

2 (bases 1 to 2364)

Isogai,T. and Yamamoto,J.

Direct Submission

Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 291-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; CDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

FEATURES

source

1. .2364

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="UTERU2002964"
/tissue_type="uterus"
/clone_lib="UTERU2"
/note="cloning vector: pME18SFL3"

BASE COUNT 420 a 768 c 776 g 400 t
ORIGIN

Query Match 75.3%; Score 418; DB 9; Length 2364;

Best Local Similarity 85.3%; Pred. No. 1.2e-46;

Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Qy 1 CACACCCACAGGACTTCAGCTGTGTGTCACCTGTCGCTGTCGACGCGCGCGCG 60

Db 1644 CACAGCACCGCGACTTCAGCGCGGTGCTCCACCTGTTGCGCTCAACAGCCCTGTCA 1703

Qy 61 GGGCGATGCGAGGACATCCGGGAGGGGACTTCCAGTGTCTCCAGCAGCGCGCGCGG 120

Db 1704 GGGCGATGCGGGGATCCGGGGGCGGACTTCCAGTGTCTCCAGCAGCGCGCGCGG 1763

Qy 121 GGGCTGGCGCGCACCTTCCGGGCGCTTCTGTGCTGTCGCGCTGCAGGACCTCTACAGCATC 180

Db 1764 GGGCTGGCGCGCACCTTCCGGGCGCTTCTGTGCTGTCGCGCTGCAGGACCTGTACAGCATC 1823

Qy 181 GTGCGCGCGCGCGACCGCACCGGGGTGCGCTGCTCAACTCAGGACCGAGGTGCTCTTC 240

Db 1824 GTGCGCGGTGCGCGACCGCGCGCGGTGCGCTCAACTCAGGACCGAGGTGCTCTTC 1883

Qy 241 CCCAGCTGGAGGCGCTTATTTCTCGGGCTCCGAGGCGCGAGTCAAGCCCGGGCGCGCATC 300

Db 1884 CCCAGCTGGAGGCGCTTGTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGCGCGCATC 1943

Qy 301 TTCTCTTTTCGCGCAGAGATGTCTGCGCAGCACCCCGCTGCGCCCGGAGAGCGGTGG 360

Db 1944 TTCTCTTTTCGCGCAGAGAGCTCTCTGAGGACCCCGCTGCGCCCGGAGAGCGGTGG 2003

Qy 361 CACGGCTCGCGACCCCGCGCGCGCGCTGACCGCAGCAGTACTGCGAGACGTGCGCGGACG 420

Db 2004 CATGGCTCGGACCCCGCGCGCGCGAGGCTGACCGAGAGCTACTGTGAGACGTGCGCGGACG 2063

Qy 421 GAGGCGCGCGCGCGCACCGGGCAGGCGTGTGCTGCTGCGCGGCGAGGCTGCTGGAGCAG 480

Db 2064 GAGGCTCTCTCGCGCGCGCGCGCGCTCTCTGCTGCTGCGGGGCGAGGCTCTCTGGGCGAG 2123

Qy 481 GAGGCGCGCGCTGCGCGCACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

Db 2124 AGTGGCGGAGTGGCATCCGCTTACATCGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2183

Qy 541 TCCTTC 546

Db 2184 GCCTCC 2189

RESULT 12

AX409531

LOCUS

Sequence 2178 from Patent WO0229103.

AX409531

AX409531

VERSION

AX409531.1

GI:21442236

KEYWORDS

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.

Gene expression profiles in liver cancer

Patent: WO 0229103-A 2178 11-APR-2002;

GENE LOGIC INC (US)

FEATURES

Location/Qualifiers

1. .3394

/organism="Homo sapiens"

/mol_type="genomic DNA"


```
CDS
<1..2055
/contid="COL18A1"
/codon_start=1
/product="collagen type XVIII alpha 1"
/db_xref="GI:562794"
/translation="GEVADGIGPGFGLPGRREGIAGPOQPKDRSGRGEKDPGKDLG
GQGLPGRPGPVVYVSEQDGSVLSVPGEGRGFAGFPAGPKNGIAGSKGLGS
PGKLGKEGPGSIFSPDGLGAPQAKGAKBEPFRPGPLGYRPGYKEIGFPGPRGR
PGMGLGKEGPGDASLFGMRGMPGPGPGPGTGVYDSNVFAESRPGPGG
LPGNGPGPKGPEVPPGPGPFDFLOKEAEMGEKDRGADGAKGGERGEGG
GGFGSSLPAGAPGPGYGPPIGPGRGESIRGQPGPGPGPGPGPGPGPGPG
GGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG
EGWLFVAEQEELVYVQNGFRKQLEARTPLPRGTNEVALQPVVQLADSNPYPR
REHPHTAPRWADILLASPGPLPQPVPGPHSSVYHCGAPRTGPPAHSHRDFQ
PVLHLVALNSPLSGMWRGIRGADFOQARAVGLAGTFRAFLSRLQDLXSVIRAD
RAAFIVNLKDELLPSPWEALFSGSEGLKPGARIFSDGKDLRHPTPQKSVWHGS
DNGRRRLTESYCTWRTEAPSATGQASSLLGRLLGQSAASCHHAYIVLCIENSFWTA
SK"
BASE COUNT 668 a 1130 c 1080 g 516 t
ORIGIN
Query Match 75.3%; Score 418; DB 9; Length 3394;
Best Local Similarity 85.3%; Pred. No. 1.1e-46;
Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 1 CACACCCACGAGCTCCAGCTGGTGCTGACCTGTGGCCCTGAACAGCCGCGAGCG 60
DB 1504 CACAGCCACGAGCTCCAGCGGGTGTCTCCAGCTGTGGCTCAACAGCCCTGTCA 1563
QY 61 GCGCGCATCGAGGATCCGCGGAGCGGACTTCCAGTGTCTTCCAGCAGCGCGCGCG 120
DB 1564 GCGCGCATCGCGGATCCGCGGCGCGGACTTCCAGTGTCTTCCAGCAGCGCGCG 1623
QY 121 GCGCTGCGCGGACCTTCCGCGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
DB 1624 GCGCTGCGCGGACCTTCCGCGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1683
QY 181 GTGCGCGCGCGACCGCGCGGCGCGCGCTGCTCAACCTCAGGACGAGGTGCTTTC 240
DB 1684 GTGCGCGCGCGACCGCGCGGCGCGCGCTGCTCAACCTCAGGACGAGGTGCTTTC 1743
QY 241 CCCAGCTGGAGGCGCTTATCTCGGCTCCGAGGCGGCGAGCTGAAGCCCGCGCGCATC 300
DB 1744 CCCAGCTGGAGGCTGTGTCTCAGGCTCTGAGGCTCCGCTGAAGCCCGCGCGCATC 1803
QY 301 TTCTCTTTCGAGCGAGAGATGCTGTGAGCACCCTGCGCGCGCGGAGAGCGGTGG 360
DB 1804 TTCTCTTTCGAGCGAGAGCTGCTGAGGACCCCACTGCGCGCGCGGAGAGCGGTGG 1863
QY 361 CACGGCTCCGACCGCGCGGCGCGCTGACCGCAGCTACTGCGAGACGTCGCGGAGC 420
DB 1864 CATGGCTCCGACCGCGCGGCGCGCTGACCGCAGCTACTGCGAGACGTCGCGGAGC 1923
QY 421 GAGGCGCGCGCGGCGGACCGCGGAGGCGCTGCTGCTGCTGCGCGGCGAGGCTGCTG 480
DB 1924 GAGGCTCCCTCGGCGGCGGCGGAGGCTCTGCTGCTGCTGCGGCGGCGGAGC 1983
QY 481 GAGGCGCGGAGTGGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 1984 AGTGCGCGGAGTGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2043
QY 541 TCCTTC 546
DB 2044 GCCTCC 2049
RESULT 15
AF018082 Homo sapiens type XVIII collagen (COL18A1) mRNA, linear PRI 18-MAR-1998
LOCUS AF018082 5408 bp mRNA linear PRI 18-MAR-1998
DEFINITION Homo sapiens type XVIII collagen (COL18A1) mRNA, alternatively
spliced, short form, complete cds.
ACCESSION AF018082
AF018082.1 GI:2920536
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 5408)
Saarela, J., Vliakopoulou, R., Rehn, M., Purmonen, S. and Pihlajaniemi, T.
Complete primary structure of two variant forms of human type XVIII
collagen and tissue-specific differences in the expression of the
corresponding transcripts
Matrix Biol. 16 (6), 319-328 (1998)
9503365
98164096
2 (bases 1 to 5408)
Saarela, J., Vliakopoulou, R., Rehn, M., Purmonen, S. and Pihlajaniemi, T.
Direct Submission
Submitted (28-JUN-1997) Dept. of Medical Biochemistry, University
of Oulu, Kajaanintie 52 A, Oulu 90220, Finland
Location/Qualifiers
1..5408
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..5408
/gene="COL18A1"
60..4070
/gene="COL18A1"
/note="alternatively spliced; short form (NCI-303)"
/product="type XVIII collagen"
/db_xref="GI:2920537"
/translation="MAPRCMPWPRRRLLDLVLLVLLGVRAASAPERISEEVL
LQLDPPPPQVTQTDSDVGLAYVFGDANSQVARIHPSLFRDLSLLHPIRAT
EGVGFALITSAQAVMLLVKLSGVQGHODISLITYEPGAGQTHTAASFLPAPVG
QWHLASVAGFVALYVDFCEFRMPLARSRLGLEFAGLFAVQAGADPDKFG
VIAELKVRDDQVSPHMLCDBEDDSDGAFDGSGLDARELLREETGAALKRLPA
PPVPTPLAGSSTEDSRSEEBQTTVASLGAOTPLGSDSVSTWDSVTPGGRVK
EGGLKQKGEVGPVGPGRAGPSPGLPGLPCVSPGLPAGPALOTVPGPGPP
GPPGKGTPTFRHDKLTITDMEGSGFDLALRPGFPVPGPGVGLPGLPGLP
QPPPGPPGPRHDKLTITDMEGSGFDLALRPGFPVPGPGVGLPGLPGLP
GVNSDVPAGLPGVPGRGPVPGPLPGPPPGPPPGPPPGPPPGPPPGPPPG
KSGKAGPAGAGSGSLAGAPGAGPPGPPPGPPPGPPPGPPPGPPPGPPG
ARSADPGKGLGLKGLDGVGLPAGKAGVADGIPGPGPLPGREGIAGPQKRGD
RGSRGKDPKGLDGVGLPAGPPPGPVVVSQDSVLSVSPGEPGPGPPGPR
AGPKNLGSKGERGSPGKGEKGPISFSDGALGPAQKAGKGPGRGPPGPR
PGYKEIGFPGPRPGRMNGLKGEKGPISFSDGALGPAQKAGKGPGRGPPG
DSNVFAESRPGPGPLPGNPGPGPKAGKEVPGPPGPPGPPGPPGPPGPPG
RGDQKGERGPGGPGGSGSLPGPPPGPGPGPGPGPGPGPGPGPGPGPGPG
PGIGYEGRQGPVPGPPPGPPPGPHRTISVPGPGPGPGPGPGPGPGPGPG
ATROAMLQVHEVPEGWLI FVAEOELVYVQNGFRKQLEARTPLPRGTNEVALQ
PVLHLVALNSPLSGMWRGIRGADFOQARAVGLAGTFRAFLSRLQDLXSVIRAP
RPTSPSHRDFQVPLHLVALNSPLSGMWRGIRGADFOQARAVGLAGTFRAFLS
SRLQDLXSVIRAPRAAFIVNLKDELLPSPWEALFSGSEGLKPGARIFSDGKDL
RHPTPQKSVWHGS DNGRRRLTESYCTWRTEAPSATGQASSLLGRLLGQSAASCHH
AVIVLCIENSFWTASK"
BASE COUNT 986 a 1920 c 1796 g 806 t
ORIGIN
Query Match 75.3%; Score 418; DB 9; Length 5408;
Best Local Similarity 85.3%; Pred. No. 9.1e-47;
Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 1 CACACCCACGAGCTTCCAGCTGGTGCTGACCTGTGGCCCTGAACAGCCGCGAGCG 60
DB 3519 CACAGCCACGCGACTTCCAGCGGCTGTCTCCAGCTGTGGCTCAACAGCCCTGTCA 3578
QY 61 GCGCGCATCGAGGATCCGCGGAGCGGACTTCCAGTGTCTTCCAGCAGCGCGCGCG 120
DB 3579 GCGCGCATCGCGGATCCGCGGAGCGGACTTCCAGTGTCTTCCAGCAGCGCGCGCG 3638
QY 121 GCGCTGCGCGGACCTTCCGCGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 12:12:16 ; Search time 213.739 Seconds
(without alignments)
7009.423 Million cell updates/sec

Title: US-09-938-391-3
Perfect score: 555
Sequence: 1 cacaccaccaggacttcca.....tgacctcttctccaagtag 555

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 19Jun03: *
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT: *
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT: *
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT: *
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT: *
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT: *
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT: *
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT: *
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT: *
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT: *
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT: *
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT: *
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT: *
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT: *
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT: *
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT: *
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT: *
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT: *
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT: *
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT: *
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: *
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT: *
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: *
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT: *
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: *
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	555	100.0	555	24	Canine endostatin
2	555	100.0	829	24	Canine pro-endosta
3	550.4	99.2	552	21	Canine angiogenesi
4	419.6	75.6	1564	24	DNA encoding human
5	418	75.3	546	21	Human endostatin g
6	418	75.3	549	20	Human endostatin D
7	418	75.3	549	21	Human angiogenesis
8	418	75.3	549	21	Human angiogenesis

9	418	75.3	549	22	AAS00867	Human gene fragmen
10	418	75.3	549	25	ABA00774	Human endostatin c
11	418	75.3	552	20	AX35375	SEQ ID 50 of WO991
12	418	75.3	552	21	AAC62023	Nucleotide sequenc
13	418	75.3	552	21	AA68203	Human endostatin c
14	418	75.3	552	22	AAC88289	Human endostatin c
15	418	75.3	552	24	ABK50685	cDNA encoding huma
16	418	75.3	555	22	AAI66529	Human vascular end
17	418	75.3	558	24	ABA99261	Human endostatin c
18	418	75.3	563	20	AZ08750	Human endostatin c
19	418	75.3	816	21	AA64013	DNA encoding a hum
20	418	75.3	900	24	ABK09977	Synthetic plasmid
21	418	75.3	968	21	AAC62025	Nucleotide sequenc
22	418	75.3	3394	18	AA78484	Human alpha-1 coll
23	418	75.3	3394	20	AX78379	Human alpha1 (XVII
24	418	75.3	3394	24	ABN95680	Gene #2178 used to
25	418	75.3	3394	24	ABN94763	Human pancreatic c
26	418	75.3	4551	24	ABN85301	Human collagen XVI
27	418	75.3	4875	24	ABQ54955	Human collagen anti
28	418	75.3	5408	20	AX77720	Human collagen 18
29	417	75.1	551	24	ABQ81193	Human endostatin c
30	415.8	74.9	641	22	AAH79104	Human endostatin e
31	415.8	74.9	641	24	AA155454	Specific tumour ce
32	415.8	74.9	641	24	AA144000	Reproductive recom
33	415.8	74.9	641	24	ABQ76079	Anticancer gene-as
34	410.8	74.0	534	21	AAA27005	Alternate human en
35	410.8	74.0	537	22	AAS00868	Human gene fragmen
36	409.4	73.8	6462	21	AAA64014	Nucleotide sequenc
37	398.4	71.8	650	24	ABK09978	Synthetic plasmid
38	394.2	71.0	552	21	AAZ50398	Human endostatin c
39	375.8	67.7	555	20	AX84635	Mouse endostatin c
40	375.8	67.7	555	20	AX79949	Endostatin coding
41	375.8	67.7	555	22	AAC88290	Murine endostatin
42	375.8	67.7	558	22	AA18701	Mouse endostatin e
43	375.8	67.7	558	24	ABQ76173	Rat endostatin enc
44	375.8	67.7	558	24	ABQ76178	Rat endostatin enc
45	375.8	67.7	565	20	AX58740	DNA encoding anti-

ALIGNMENTS

RESULT 1
AAL46063
ID AAL46063 standard; cDNA; 555 BP.

XX AAL46063;

AC AAL46063;

XX 19-JUL-2002 (first entry)

XX Canine endostatin coding sequence.

XX Dog; pro-endostatin; endostatin; angiogenesis; cancer; gene therapy;
KW psoriasis; rheumatoid arthritis; retinopathy; macular degeneration;
KW corneal graft rejection; neovascular glaucoma; retrolental fibroplasia;
KW rheositis; Osler-Webber Syndrome; myocardial angiogenesis;
KW plaque neovascularisation; telangiectasia; haemophilic joints;
KW angiofibroma; wound granulation; coronary collateral;
KW ischaemic limb angiogenesis; diabetic neovascularisation; fracture;
KW cytostatic; antiarthritic; antipsoriatic; antidiabetic; ophthalmological;
KW gynaecological; gene; ss.

OS Canis familiaris.

XX Location/Qualifiers
Key 1..555
CDS /*tag= a
/product= "endostatin"
/partial
/note= "no start codon"

PN EP1191036-A2.

RESULT 3	
AAZ51309	AAZ51309 standard; DNA; 552 BP.
XX	
XX	
XX	AAZ51309;
XX	
XX	06-JUN-2000 (first entry)
XX	
XX	Canine angiogenesis inhibitor, endostatin DNA.
XX	
KW	Canine; immunoglobulin Fc fragment; endostatin; immunofusin;
KW	angiogenesis; inhibitor; cytostatic; antiinematoid; antiarthritic;
KW	antipsoriatic; antidiabetic; ophthalmological; immunosuppressant;
KW	vasotropic; vulnery; treatment; antiarteriosclerosis; tumour;
KW	metastasis; atherosclerosis; psoriasis; rheumatoid arthritis;
KW	ocular angiogenic disease; diabetic retinopathy; macular degeneration;
KW	myocardial angiogenesis; plaque neovascularisation; telangiectasia;
KW	wound granulation; keloid scar; gene therapy; ds.
XX	
XX	Canis familiaris.
XX	
XX	
KW	Location/Qualifiers
FH	CDS
FT	1..552
FT	/*tag= a
FT	/product= "Endostatin"
FT	/note= "Does not include stop codon"
FT	/partial
XX	
XX	WO200011033-A2.
PN	
XX	

Query Match	99.28;	Score 550.4;	DB 21;	Length 552;
Best Local Similarity	99.84;	Pred. No. 1.5e-84;		
Matches 551;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0
Qy	1	CACACCCACAGAGACTTCCAGCTGGTGCTGCACCTGGTGGCCCTGAAACAGCCCGCAGCGC	60	
Db	1	CACACCCACAGAGACTTCCAGCCGGTCTGCACCTGGTGGCCCTGAAACAGCCCGCAGCGC	60	
Qy	61	GGCGGATGAGAGCATCCGGGAGCGGACTTCCAGTGTCTTCAGCAGCGCGCGCCGCG	120	
Db	61	GGCGGATGAGAGCATCCGGGAGCGGACTTCCAGTGTCTTCAGCAGCGCGCGCGCG	120	
Qy	121	GGGCTGCCCGGCACTTCCCGGGCCCTTCTGTCTGCTCGGGCTGCAGGACCTCTACAGCATC	180	
Db	121	GGGCTGCCCGGCACTTCCCGGGCCCTTCTGTCTGCTCGGGCTGCAGGACCTCTACAGCATC	180	
Qy	181	GTGCGCGCGCCGACCGCACCGGGGTGCCCGTCTGCTCAACTTCAGGAGCAGGTGCTCTTC	240	
Db	181	GTGCGCGCGCCGACCGCACCGGGGTGCCCGTCTGCTCAACTTCAGGAGCAGGTGCTCTTC	240	
Qy	241	CCAGCTGGAGGCGCTTATTTCTGGGCTCCGAGGGCCAGCTTGAAGCCCGGGGCCCGCATC	300	
Db	241	CCAGCTGGAGGCGCTTATTTCTGGGCTCCGAGGGCCAGCTTGAAGCCCGGGGCCCGCATC	300	
Qy	301	TTCTCTTTTCAGCGCAGAGATGTCTCTGACGACCCCGGCTTGCCGCCCGGAAGAGCGTGTTGG	360	
Db	301	TTCTCTTTTCAGCGCAGAGATGTCTCTGACGACCCCGGCTTGCCGCCCGGAAGAGCGTGTTGG	360	
Qy	361	CAGGCTCCGACACCCAGCGGGCGCGCTGTACCGACAGCTACTGTGGAGACGTGGCGGACG	420	
Db	361	CAGGCTCCGACACCCAGCGGGCGCGCTGTACCGACAGCTACTGTGGAGACGTGGCGGACG	420	

XX. 30-OCT-1998; 98US-0106343.
PR 20-MAY-1999; 99US-0315689.
XX (CHIL-) CHILDRENS MEDICAL CENT.
PA
XX O'Reilly MS, Folkman MJ;
PI WPI; 2000-365617/31.
XX P-PSDB; AAY94323.
DR
XX Novel endostatin capable of inhibiting endothelial cell proliferation
PT and angiogenesis, useful for treating angiogenesis-dependent cancers
PT and as birth control agents
XX
PS Claim 10; Page 39; 68pp; English.
XX The present sequence encodes an endostatin protein which is the carboxy
CC terminal protein of human collagen XVIII. Recombinant mouse endostatin
CC (20 mg/kg) was administered subcutaneously to mice implanted with Lewis
CC lung carcinomas. There was tumour mass regression non-detectable levels
CC after 12 days of therapy due to the angiogenesis inhibitory activity of
CC endostatin. Thus the protein is useful for treatment of angiogenesis-
CC dependent cancers. The polynucleotide and polypeptide sequences of this
CC endostatin are useful for treating and diagnosis of tumours, ocular
CC angiogenic diseases, Osler-Weber syndrome, myocardial angiogenesis,
CC plaque neovascularisation, telangiectasia, haemophilic joints,
CC angiofibroma and wound granulation, for treatment of diseases related to
CC excessive or abnormal stimulation of endothelial cells e.g. intestinal
CC adhesions, atherosclerosis, scleroderma. The protein may also be useful
CC as a birth control agent by reducing or preventing uterine
CC vascularisation. The gene for endostatin may be isolated from cells or
CC tissue that express high levels of endostatin, eg. tumour cells, by
CC generating cDNA from mRNA using reverse transcriptase and then amplifying
CC the DNA sequence.
XX
SQ Sequence 546 BP; 80 A; 196 C; 177 G; 93 T; 0 other;
Query Match 75.3%; Score 418; DB 21; Length 546;
Best Local Similarity 85.3%; Pred. No. 3.4e-62;
Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 1 CACACCCACGAGACTTCCAGCTGTGTGCTGCACCTGGTGGCCCTGAACAGCCGCGAGCGG 60
DB 1 CACAGCCACGCGACTTCCAGCGGTGCTCCACCTGGTGGCTCAACAGCCCTGTCA 60
QY 61 GCGGCGATCGAGGATATCCGGGAGCGGACTTCCAGTGTTCAGAGCGCGCGCGCG 120
DB 61 GCGGCGATCGGGGATCCGCGGGCCGACTTCCAGTGTTCAGAGCGCGCGCGCG 120
QY 121 GGGCTGGCGGCACCTTCCGGGCTTCTGCTGCTGGCGCTGCGAGCCTTACAGCATC 180
DB 121 GGGCTGGCGGCACCTTCCGGGCTTCTGCTGCTGGCGCTGCGAGCCTTACAGCATC 180
QY 181 GTGCGCGCGCGACCGACCGGGGTGCGCGTCTGCTCAACTCAGGACGAGGTGCTTCTTC 240
DB 181 GTGCGCGGTGCGACCGCGCGAGCGGTGCCATGCTCAACTCAGGACGAGGTGCTTTC 240
QY 241 CCCAGCTGGAGGCGCTTATTTCTCGGGCTCGAGGGCCAGTGAAGCCCGGGGCGCGCATC 300
DB 241 CCCAGCTGGAGGCGCTTATTTCTCGGGCTCGAGGGTCCGCTGAAGCCCGGGGCGCGCATC 300
QY 301 TTCTCTTTTCGAGGAGATGTCCTGAGCAGCCCGCTGCGCCCGGAGAGCGGTGCG 360
DB 301 TTCTCTTTTCGAGGAGGAGCTCTGTTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGGCGCGCATC 360
QY 361 CAGGCTTCGAGCCCGAGCGGGCCCGCTGACCGACAGCTACTTGCAGACGTGGCGGAGC 420
DB 361 CATGGCTCGGACCCCAACCGGGCGAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGAGC 420
QY 421 GAGGCCCCCGCGCCACCGGGCAGGCGTCTCGCTGCTCGCGGCGAGGCTGTGAGGAGCAG 480
DB 421 GAGGCTCCCTCGGCCACCGGCCAGGCGCTCTCTGCTGCTGGGGGCGAGGCTCTCTGGGGCAG 480

QY 481 GAGCGCGGAGCTCCGCGCACCCCTTCTGCTGTGGTGTCTGTGATCGAGAACAGCGCTCATGACC 540
DB 481 AGTGCGCGGAGCTGCCATCAGCGCTTACATCGTGTCTGTGATCGAGAACAGCGCTCATGACT 540
QY 541 TCCTTC 546
DB 541 GCTCC 546
RESULT 6
AAX77719
ID AAX77719 standard; DNA; 549 BP.
XX
AC AAX77719;
XX
DT 10-AUG-1999 (first entry)
XX Human endostatin DNA coding region fragment.
DE
XX Plasmidogen; human; angiotensin; endostatin; gene therapy; vector;
XX anti-angiogenic; attenuation; cytostatic; anti-diabetic; ophthalmology;
KW tumour growth; solid tumour; diabetic retinopathy; retina; ss.
KW
XX Homo sapiens.
OS
XX WO9926480-A1.
PN
XX 03-JUN-1999.
PD
XX 20-NOV-1998; 98WO-US24950.
PF
XX 20-NOV-1997; 97US-0975424.
PR
XX (GENE-) GENETIX PHARM INC.
PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Bachelot T, Leboulch P, Pawliuk RJ;
XX
DR WPI; 1999-357696/30.
XX P-PSDB; AAY08693.
XX
XX Anti-angiogenic gene therapy vectors
PT
PS Disclosure; Page 74; 83pp; English.
XX
CC This invention describes a novel viral gene therapy vector comprising a
CC nucleic acid molecule encoding an anti-angiogenic polypeptide chosen
CC from human or murine angiotensin, human or murine endostatin and
CC angiogenesis-inhibiting fusions and fragments, where the viral vector is
CC sufficiently attenuated for use in human gene therapy. The products of
CC the invention have anti-angiogenic, cytostatic, anti-diabetic and
CC ophthalmological activity. The vector is used in gene therapy for
CC inhibiting tumour growth in humans harbouring a solid tumour. The vector
CC expresses an anti-angiogenic polypeptide. An additional use comprises
CC treatment of diabetic retinopathy, where the anti-angiogenic polypeptide
CC inhibits angiogenesis in the vicinity of the retina. The vector is
CC administered to cells ex vivo and then administered to the patient.
XX
SQ Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;
Query Match 75.3%; Score 418; DB 20; Length 549;
Best Local Similarity 85.3%; Pred. No. 3.4e-62;
Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 1 CACACCCACGAGACTTCCAGCTGTGTGCTGCACCTGGTGGCCCTGAACAGCCGCGAGCGG 60
DB 1 CACAGCCACGCGACTTCCAGCGGTGCTCCACCTGGTGGCTCAACAGCCCTGTCA 60
QY 61 GCGGCGATCGAGGATATCCGGGAGCGGACTTCCAGTGTTCAGAGCGCGCGCGCG 120
DB 61 GCGGCGATCGGGGATCCGCGGGCCGACTTCCAGTGTTCAGAGCGCGCGCGCG 120

QY 121 GGCTGGCGGCGACCTTCGGGCTTCTCTGTCGCGGCTGAGGACCTCTACAGCATC 180
 Db |||||
 QY 121 GGCTGGCGGCGACCTTCGGGCTTCTCTGTCGCGGCTGAGGACCTGTACAGCATC 180
 Db |||||
 QY 181 GTGCGCGCGCGACCGCACCGGGGTGCGGCTCAACCTTAGGACAGGTGCTTTC 240
 Db |||||
 QY 181 GTGCGCGGTGCGGCGCGCGCGGTGCGGCTCAACCTTAGGACAGGTGCTTTC 240
 Db |||||
 QY 241 CCAGCTGGGAGGCTTATCTCGGCTCCGAGGCGCAGTGAAGCCCGGGCCCGCATC 300
 Db |||||
 QY 241 CCAGCTGGGAGGCTTCTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGCACGCATC 300
 Db |||||
 QY 301 TTCTCTTTTCAGCGCAGAGATGCTTCGACGACCCCGGCTGCGGCGGAGGCGGTGG 360
 Db |||||
 QY 301 TTCTCTTTTCAGCGCAGAGATGCTTCGACGACCCCGGCTGCGGCGGAGGCGGTGG 360
 Db |||||
 QY 361 CACGGCTCCGACCCGAGCGGCGCGCTGACGACAGCTACTGCGAGACGTGGCGGAGC 420
 Db |||||
 QY 361 CATGGCTCGGACCCCAACGGGCGCAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGAGC 420
 Db |||||
 QY 421 GAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
 Db |||||
 QY 421 GAGGCTCCCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
 Db |||||
 QY 481 GAGGCGCGGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
 Db |||||
 QY 481 AGTGCGCGGAGCTGCCATCAGCGCTACATCGTGTCTGCTGATTGAGAACAGCTTCATGACT 540
 Db |||||
 QY 541 TCCTTC 546
 Db |||||
 QY 541 GCCTCC 546

RESULT 7

AAA29884
 ID AAA29884 standard; cDNA; 549 BP.
 AC AAA29884;
 DT 22-AUG-2000 (first entry)
 DE Human angiogenesis inhibiting factor 1 encoding cDNA.
 XX Human; angiogenesis inhibiting factor 1; IAF-1; tumour; antibody;
 XX abnormal vessel disease; ss.
 OS Homo sapiens.
 XX CN1244536-A.
 PD 16-FEB-2000.
 PF 10-AUG-1998; 98CN-0117150.
 PR 10-AUG-1998; 98CN-0117150.
 PA (ONCO-) INST ONCOLOGY UNDER TUMOR HOSPITAL CHINE.
 XX Yang Z, Guo W;
 XX WPI; 2000-388168/34.
 DR P-PSDB; AAY90771.
 XX Angiogenesis inhibiting factor 1 and its derivative useful for treating
 PT tumors -
 XX Example 1; Fig 5; 41pb; Chinese.
 XX The present sequence encodes an angiogenesis inhibiting factor (1),
 CC designated IAF-1. The present invention also describes: (1) preparation
 CC of (1) and its derivative; (2) an IAF binding acceptor and its
 CC preparation; and (3) an IAF antibody. (1) is useful for preparing new
 CC biological preparations for effectively treating various tumours and

CC abnormal-vessel diseases. The IAF antibody is preferably a polyclonal
 CC antibody, mosaic antibody, single stranded antibody and human originated
 CC antibody.
 XX

SQ Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;

Query Match 75.3%; Score 418; DB 21; Length 549;

Best Local Similarity 85.3%; Pred. No. 3.4e-62;

Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 1 CACACCCACAGGACTTCAGTGTGTGACCTGGTGGCCCTGAAAGCCCGGAGCGG 60
 Db |||||
 QY 1 CACAGCCACCGGACTTCAGCCGCTGCTCCACCTGGTGGCTCAACAGCCCCGTGCA 60
 Db |||||
 QY 61 GCGCGCATGCGGAGCATCCGGGAGCGGACTTCCAGTGTCTCCAGCAGCGCGCCCGG 120
 Db |||||
 QY 61 GCGCGCATGCGGAGCATCCGGGAGCGGACTTCCAGTGTCTCCAGCAGCGCGCCCGG 120
 Db |||||
 QY 121 GGGCTGGCGGCGGACCTTCGCGGCTTCTGTGTCGCGGCTGACGAGCTCTACAGCATC 180
 Db |||||
 QY 121 GGGCTGGCGGCGGACCTTCGCGGCTTCTGTCTCGCGCTGCAGGACCTGTACAGCATC 180
 Db |||||
 QY 181 GTGCGCGCGCGGACCGGACCGGCGGCTGCTGCTCAACCTCAGGACAGGTGCTTTC 240
 Db |||||
 QY 181 GTGCGCGCGGCGGACCGGCGGCGGCTGCTGCTCAACCTCAGGACAGGTGCTTTC 240
 Db |||||
 QY 241 CCCAGCTGGGAGGCTTATCTCGGCGCTCCGAGGCGGCGGCTGAAAGCCCGGCGGCGG 300
 Db |||||
 QY 241 CCCAGCTGGGAGGCTTCTGTCTCAGGCTCTGAGGCTCGCTGAAAGCCCGGCGGCGG 300
 Db |||||
 QY 301 TTCTCTTTTCAGCGCAGAGATGCTTCGACGACCCCGGCTGCGGCGGAGGCGGTGG 360
 Db |||||
 QY 301 TTCTCTTTTCAGCGCAGAGATGCTTCGAGGCGGCGGCGGCGGCGGCGGCGGCGG 360
 Db |||||
 QY 361 CACGGCTCCGACCCGAGCGGCGGCGGCTGACGAGCTACTGCGAGACGTGGCGGAGC 420
 Db |||||
 QY 361 CATGGCTCGGACCCCAACGGGCGCAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGAGC 420
 Db |||||
 QY 421 GAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
 Db |||||
 QY 421 GAGGCTCCCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
 Db |||||
 QY 481 GAGGCGCGGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
 Db |||||
 QY 481 AGTGCGCGGAGCTGCCATCAGCGCTACATCGTGTCTGCTGATTGAGAACAGCTTCATGACT 540
 Db |||||
 QY 541 TCCTTC 546
 Db |||||
 QY 541 GCCTCC 546

RESULT 8

AAZ51291
 ID AAZ51291 standard; cDNA; 549 BP.

AC AAZ51291;

DT 06-JUN-2000 (first entry)

DE Human angiogenesis inhibitor, endostatin cDNA.

KW Human; immunoglobulin gamma Fc fragment; endostatin; immunofusin;
 KW angiogenesis; inhibitor; cytostatic; antirheumatoid; antiarthritic;
 KW antiproliferative; antidiabetic; ophthalmological; immunosuppressant;
 KW vasotropic; vulnery; treatment; antiarteriosclerosis; tumour;
 KW metastasis; atherosclerosis; psoriasis; rheumatoid arthritis;
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
 KW wound granulation; keloid scar; gene therapy; ss.

OS Homo sapiens.

XX Key

Location/Qualifiers

FH

FT CDS 1..549
 FT /tag= a
 FT /product= "Endostatin"
 FT /note= "Does not include stop codon"
 FT /partial
 FT XX WO200011033-A2.
 PN 02-MAR-2000.
 PD 25-AUG-1999; 99WO-US19329.
 PF 25-AUG-1998; 98US-0097883.
 PR (LEXI-) LEXINGEN PHARM CORP.
 PA Lo K, Li Y, Gillies SD;
 PI WPI; 2000-237616/20.
 DR P-PSDB; AAY70252.
 XX Novel fusion protein of angiostatin or endostatin and an immunoglobulin
 PT FC region, useful for treating conditions mediated by angiogenesis,
 PT such as rheumatoid arthritis, tumors and macular degeneration -
 XX
 PS Claim 12; Pages 40-41; 69pp; English.
 CC The patent discloses a DNA molecule encoding a fusion protein comprising
 CC a signal sequence, an immunoglobulin Fc region, and an angiogenesis
 CC inhibitor selected from angiostatin, endostatin, a plasminogen fragment
 CC having angiostatin activity, a collagen XVIII fragment having endostatin
 CC activity, or combinations of them. The fusion protein (immunofusin) is
 CC used to inhibit angiogenesis and to treat diseases or conditions mediated
 CC by angiogenesis. Conditions that may be treated include solid tumours,
 CC blood born tumours, tumour metastasis, benign tumours including
 CC haemangiomas, acoustic neuromas, neurofibromas, trachomas and pyrogenic
 CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases
 CC e.g. diabetic retinopathy, retinopathy of prematurity, macular
 CC degeneration, corneal graft rejection, neovascular glaucoma, retrolental
 CC fibroplasia, rubeosis and Osler-Weber syndrome; myocardial angiogenesis,
 CC plaque neovascularisation, telangiectasia, haemophilic joints,
 CC angiofibroma, wound granulation, and excessive or abnormal stimulation of
 CC endothelial cells, intestinal cells, atherosclerosis, scleroderma and
 CC hypertrophic scars, i.e. keloid scars. The DNA constructs may be used
 CC in gene therapy. The present sequence is a cDNA encoding human
 CC endostatin used in the construction of immunofusin containing human
 CC immunoglobulin gamma (IgG) Fc fragment.
 CC Note: This sequence is stated in claim 12 as being amino acid sequence
 CC of plasminogen fragment, however, the rest of the specification
 CC refers to this sequence as being nucleotide sequence of human endostatin.
 XX
 SQ Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;

Query Match 75.3%; Score 418; DB 21; Length 549;
 Best Local Similarity 85.3%; Pred. No. 3.4e-62;
 Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 1 CACACCCACGAGACTTCCAGTGTGCTGTCACCTGGTGGCTGACAGCCGCGCGG 60
 DB 1 CACAGCCACCGGACTTCCAGCGGGTGTCCACCTGGTGGCTCAACAGCCCTGTCA 60
 QY 61 GCGGGCATGCGAGGATCCGGGAGCGGACTTCCAGTGTCTCCAGCAGCGCGCGCGG 120
 DB 61 GCGGGCATGCGGGGATCCGCGGGGCGGACTTCCAGTGTCTCCAGCAGCGCGCGG 120
 QY 121 GGGCTGGCGCGCACTTCCGGGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
 DB 121 GGGCTGGCGCGCACTTCCGGGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
 QY 181 GTGCGCGCGCGCACTTCCGGGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
 DB 181 GTGCGCGCGCGCACTTCCGGGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240

QY 241 CCCAGCTGGAGGCTTATTCTCGGGCTCCGAGGGCCAGCTGAAGCCCGGGCCCGCATC 300
 DB 241 CCCAGCTGGAGGCTTATTCTCGGGCTCCGAGGGCTCGCTGAGGGTCCGCTGAAGCCCGGGCCCGCATC 300
 QY 301 TTCTCTTTGACGGGACAGAGATGTCTGACGACACCGGCTGCGCCCGGAGAGCGGTGTGG 360
 DB 301 TTCTCTTTGACGGGACAGAGATGTCTGAGGGACCCCACTTGGCCCGGAGAGCGGTGTGG 360
 QY 361 CACGGCTCCGACCCCGGCGGCGCTGACCCGACAGCTACTGCGAGACGTGGCGGACG 420
 DB 361 CATGGCTCGGACCCCAACGGGGGAGGCTGACCCGAGAGCTACTGTGAGACGTGGCGGACG 420
 QY 421 GAGGCCCGGGGCGCACCGGGCAGCGCTGCTGCTGTGGCGGGCAGGCTGTGAGACGAG 480
 DB 421 GAGGCTCCCTCGGCGCACCGGGCAGGCGCTCTCGCTGTGGGGGCGAGGCTCTCTGGGGCAG 480
 QY 481 GAGGCGCGGAGCTGCGGCGCACGCGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
 DB 481 AGTGCGCGGAGCTGCCATCAGCGCTACATCGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
 QY 541 TCCTTC 546
 DB 541 GCCTCC 546

RESULT 9
 AAS00867
 ID AAS00867 standard; DNA; 549 BP.
 AC AAS00867;
 DX 04-JUL-2001 (first entry)
 XX Human gene fragment encoding Endostatin(TM) protein.
 DE Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;
 KW blood borne tumour; leukaemia; tumour metastasis; benign tumour;
 KW haemangioma; acoustic neuroma; neurofibroma; trachoma; rubeosis;
 KW pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
 KW retinopathy of prematurity; macular corneal graft rejection;
 KW neovascular glaucoma; retrolental fibroplasia; Osler-Weber Syndrome;
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
 KW haemophilic joint; angiofibroma; wound granulation; ds.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT 1..540
 FT CDS /tag= a
 FT /product= "Endostatin(TM) C-terminus minus 3"
 FT /partial
 FT /note= "Variant produced during fermentation
 FT reaction of Pichia pastoris harbouring an expression
 FT plasmid containing the present sequence"
 FT 1..543
 FT CDS /tag= b
 FT /product= "Endostatin(TM) C-terminus minus 2"
 FT /partial
 FT /note= "Variant produced during fermentation
 FT reaction of Pichia pastoris harbouring an expression
 FT plasmid containing the present sequence"
 FT 1..546
 FT CDS /tag= c
 FT /product= "Endostatin(TM) C-terminus minus 1"
 FT /partial
 FT /note= "Variant produced during fermentation
 FT reaction of Pichia pastoris harbouring an expression
 FT plasmid containing the present sequence"
 FT 1..549
 FT CDS /tag= d
 FT /product= "Endostatin(TM)"
 FT /partial

FT /note= "None of the above CDSs have start or stop codons"
 PN WO200119989-A2.
 XX 22-MAR-2001.
 XX 14-SEP-2000; 2000WO-US25166.
 XX 14-SEP-1999; 99US-0153698.
 XX (ENTR-) EXTREMED INC.
 XX Liang H, Sim KL, Chang-Murad A, Zhou X, Madisen J, Boerner RJ;
 PI Bermejo LL, Mistry FR, Shepard SR, Schrimsher JL;
 XX WPI; 2001-244802/25.
 DR P-PSDB; AAU00896, AAU00897, AAU00898, AAU00899.
 XX Producing Endostatin protein for treating angiogenesis mediated
 PT diseases such as solid tumours, comprises recombinantly producing the
 PT protein using an expression system, and recovering and purifying the
 PT protein -
 XX Claim 6; Page 29; 67pp; English.
 XX The sequence encodes Human Endostatin(TM). The new method of the
 CC invention is useful for producing, recovering and purifying Endostatin
 CC (TM) from biological sources, such as biological fluids, tissues, cells,
 CC culture media, and fermentation media. Endostatin(TM) is useful for
 CC treating angiogenesis mediated diseases such as solid tumours, blood
 CC borne tumours, leukaemias, tumour metastases, benign tumours, e.g.
 CC haemangioma, acoustic neuromas, neurofibromas, trachomas, and pyogenic
 CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases,
 CC e.g., diabetic retinopathy, retinopathy of prematurity, macular
 CC degeneration, corneal graft rejection, neovascular glaucoma, colon
 CC cancer, retrolental fibroplasia, rubeosis, Osler-Webber Syndrome,
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,
 CC haemophilic joints, angiofibroma, and wound granulation. Endostatin(TM)
 CC is also useful for treating disease of excessive or abnormal stimulation
 CC of endothelial cells such as intestinal adhesions, atherosclerosis, and
 CC scleroderma and hypertrophic scars. Higher yields of more purified, and
 CC biologically active Endostatin(TM) are obtained by the new method.
 CC Endostatin(TM) can be stored in buffers for extended periods of time, and
 CC also subjected to lyophilisation, while preserving biological activity.
 CC Centrifugation of broth from fermentation steps in production is avoided,
 CC preventing unwanted potential cellular lysis and contamination with
 CC additional proteins, pigments, enzymes and other cellular chemicals and
 CC debris.
 XX SQ Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;
 Query Match 75.3%; Score 418; DB 22; Length 549;
 Best Local Similarity 85.3%; Pred. No. 3.4e-62;
 Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
 QY 1 CACACCCACAGGACTTCCAGTGGTGTGCTGACCTGGTGGCTGAACAGCCGCGAGCG 60
 DB 1 CACAGCCACCGACTTCCAGCGGTGTCTCACCTGTGGCTCAACAGCCCTGTCTA 60
 QY 61 GCGCGCATGCGAGGCATCCCGGGAGCGGACTTCCAGTGTCTTCAGAGCGCGCGCG 120
 DB 61 GCGCGCATGCGGGGACTCCGCGGGGCGGACTTCCAGTGTCTTCAGAGCGCGCGGTG 120
 QY 121 GCGCTGCGCGGACCTTCCCGGGCTTCTGTCTGCGGGCTGAGGACCTTACAGCATC 180
 DB 121 GCGCTGCGCGGACCTTCCCGGGCTTCTGTCTGCGGGCTGAGGACCTTACAGCATC 180
 QY 181 GTGCGCGCGCGGACCGACCGCGGGTGCCTGTGTAACCTCAGGACGAGGTCTTTC 240
 DB 181 GTGCGCGCGTCCGACCGCGGAGCGGTGCTCATCTCAACCTCAAGACGAGTCTGTTT 240
 QY 241 CCCAGCTGGAGGCTTATTCTCGGGCTCCGAGGCGCAGCTGAAGCCCGGGCCCGCATC 300
 DB 241 CCCAGCTGGAGGCTTATTCTCGGGCTCCGAGGCGCAGCTGAAGCCCGGGCCCGCATC 300

DB 241 CCCAGCTGGAGGCTTCTGTCTTCTAGGCTCCGCTGAGCCCGGGGCGCAGCATC 300
 QY 301 TTCTCTTTTCGCGGAGAGATGTCTGAGCAGCCCGCTGCGGAGAGCGTGTGG 360
 DB 301 TTCTCTTTTCGCGGAGAGATGTCTGAGCAGCCCGCTGCGGAGAGCGTGTGG 360
 QY 361 CACGGCTCCGACCCCGGCGGCGCTGACCGAGCTACTGCGAGACGTGGGGGACG 420
 DB 361 CATGGCTCGGACCCCAACGGGCGGAGGCTGACCGAGAGCTACTGAGACGTGGCGGACG 420
 QY 421 GAGGCGCGGCGGCGGCGGCGGAGGCGGCTGCTGCTGCGGCGGCGGAGGCTGCTGAGACG 480
 DB 421 GAGGCTCCCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
 QY 481 GAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
 DB 481 AGTGGCGGAGCTGCGCATCAGCCCTACATCGTGTCTGCTGCTGAGACAGCTTCATGACT 540
 QY 541 TCCTTC 546
 DB 541 GCCTCC 546
 RESULT 10
 ABA00774
 ID ABA00774 standard; CDNA; 549 BP.
 XX AC ABA00774;
 XX 18-MAR-2003 (first entry)
 XX Human endostatin coding sequence.
 XX Gene; human; plasminogen; angiotensin; neovascularisation;
 KW kringlike domain; cell proliferation; viral vector;
 KW replication-defective; cancer; tumour; ss.
 XX Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..549
 FT /*tag= a
 FT /product= "Endostatin"
 XX WO200288173-A2.
 XX PD 07-NOV-2002.
 XX 29-APR-2002; 2002WO-US13461.
 XX 30-APR-2001; 2001US-287673P.
 XX 05-APR-2002; 2002US-370634P.
 XX (CELL-) CELL GENESYS INC.
 XX Chang B, Wu WW, Macarthur J, Patel S, Jooss K, Mendez M;
 WPI; 2003-129131/12.
 P-PSDB; AAG79753.
 XX New recombinant viral vector expressing human angiotensin useful for
 PT inhibiting angiogenesis in a mammalian subject with cancer or tumor -
 XX Example 4; Page 80-82; 83pp; English.
 XX This sequence encodes endostatin. Endostatin is a 20 kD C-terminal
 CC fragment of collagen XVII that inhibits angiogenesis. The endostatin
 CC coding sequence may be used in the recombinant viral vector of the
 CC invention for obtaining angiotensin activity. The vector comprises a
 CC promoter capable of expressing human angiotensin operably linked to a
 CC structural gene encoding one or more domains of human angiotensin.
 CC The vector, which may be a replication-defective viral vector, is useful
 CC for inhibiting angiogenesis in a mammal, especially with cancer or a


```

QY 541 TCCTTC 546
DB 541 GCCTCC 546

RESULT 12
AAC62023
ID AAC62023 standard; DNA; 552 BP.
XX
AC AAC62023;
XX
DT 06-MAR-2001 (first entry)
XX
DE Nucleotide sequence of human endostatin encoded by plasmid pMALCH#15.
XX
DE Streptomyces sp. strain C5; SnpA; S. venezuelae; alpha-amylase;
KW endostatin; cancer; tumour growth; angiogenesis; ss.
XX
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..552
FT /*tag= a
FT /product= "endostatin"
XX
XX WO200060945-A1.
XX
PD 19-OCT-2000.
XX
XX 12-APR-2000; 2000WO-US09747.
XX
XX 13-APR-1999; 99US-0129084.
XX
XX (MERI ) MERCK & CO INC.
XX
XX Desanti CL, Strohl WR;
XX
XX WPI; 2000-686970/67.
XX
XX P-PSDB; AAB30493.
XX
XX Preparation of soluble recombinant endostatin involves transforming
PT Streptomyces host with expression vector comprising nucleotide
PT sequence encoding endostatin operably linked to linker and leader
PT peptide
XX
XX Example 1; Fig 6; 57pp; English.
XX
XX The present sequence encodes human endostatin. The protein is expressed
CC in Streptomyces. Leader sequences of Streptomyces sp. strain C5
CC SnpA and S. venezuelae alpha-amylase proteins are linked to the
CC N-terminal of endostatin. This ensures that endostatin protein is
CC produced as a secreted, soluble protein which needs no refolding, is
CC stable in the fermentation broth and is produced in large quantities.
CC The method is used for preparing soluble recombinant human, murine or
CC primate endostatin, which is useful in the treatment of cancer,
CC inhibition of tumour growth, inhibition of angiogenesis, isolation of
CC receptors for endostatin and for identification of anti-angiogenic
CC compounds in assays. The endostatin protein is produced as a secreted,
CC soluble protein which needs no refolding, is stable in the fermentation
CC broth and is produced in large quantities. Streptomyces are amenable
CC for cultivation in large fermentations allowing for large quantities of
CC soluble endostatin to be produced.
XX
XX Sequence 552 BP; 83 A; 196 C; 179 G; 94 T; 0 other;
XX
Query Match 75.3%; Score 418; DB 21; Length 552;
Best Local Similarity 85.3%; Pred. No. 3.4e-62;
Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
XX
QY 1 CACACCCACGAGCTCCAGCTGGTCTGCACCTGGTGGCCCTGACAGCCCGGAGCCG 60
DB 1 CACAGCCACCGGACTTCAGCGGGTGTCTCCACTGTGGCTCAACAGCCCCCTGTCA 60

```

```

61 GCGCGCATCGCAGGATCCGGGAGCGGAGCTTCCAGTGTCTCCAGACGCGCGCGCG 120
61 GCGCGCATCGGGGCGATCCGGGGCGGACTTCCAGTGTCTCCAGCAGCGCGGGCGGTG 120
121 GGGCTGGCGGACCTTTCGGGCTTCTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180
121 GGGCTGGCGGACCTTTCGGGCTTCTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180
181 GTGCGCGCGGCGACCGGACCGGCGTGGCGGTGCGTCAACCTCAGGACGAGGTGCTTTC 240
181 GTGCGCGGTGCGACCGGCGCGGCGGCGTGGCGGCTGCAACCTCAGGACGAGGTGCTTTC 240
241 CCCAGCTGGAGGCTTATTTCTGGGCTCCGAGGCGCAGCTGAAGCCCGGGGCCGCATC 300
241 CCCAGCTGGAGGCTTCTGTCTCAGGCTCTGAGGCTCGGCTGAAGCCCGGGGCCGCATC 300
301 TTCTCTTTCGAGCGGAGAGATGTCTGAGCAGCACCCTGCGCCCGGGAAGAGCGTGTGG 360
301 TTCTCTTTCGAGCGGAGAGATGTCTGAGGCGCACCCTGCGCCCGGGAAGAGCGTGTGG 360
361 CACGCTCCGACCCCGAGCGGCGCGCTGACCCGACAGCTACTGCGAGACGTGGCGGACG 420
361 CATGCTCGGACCCCAACCGGCGCGGCTGACCGAGAGTACTGTGAGACGTGGCGGACG 420
421 GAGGCTCCCTCGGCGCAGCGGCGCAGGCTCTGCTGCTGGCGGCGAGGCTGTGGAGCAG 480
421 GAGGCTCCCTCGGCGCAGCGGCGCAGGCTCTGCTGCTGGCGGCGAGGCTCTGCGGCGAG 480
481 GAGGCGCGGAGCTGCGGCGCAGGCTCTGCTGCTGGCGGCGAGGCTGTGGAGCAG 540
481 AGTGGCGGAGTGCATCACGCTTACATCGTGTGCTGTGATGAGAACAGCTTCATGACT 540
541 TCCTTC 546
541 GCCTCC 546

RESULT 13
AAA68203
ID AAA68203 standard; cDNA; 552 BP.
XX
AC AAA68203;
XX
DT 27-OCT-2000 (first entry)
XX
DE Human endostatin cDNA sequence.
XX
KW Angiogenesis-inhibiting protein receptor; angiogenesis; angiostatin;
KW endostatin; plasminogen; laminin; treatment; wound healing; solid tumour;
KW psoriasis; scleroderma; myocardial angiogenesis; Crohn's disease;
KW cerebral collateral; arteriovenous malformation; rubeosis; cancer;
KW diabetic retinopathy; arthritis; wound healing; peptic ulcer;
KW Helicobacter related disease; fracture; cat scratch fever; ss.
XX
OS Homo sapiens.
XX
PN WO200032631-A2.
XX
XX 08-JUN-2000.
XX
XX 06-DEC-1999; 99WO-US28897.
XX
XX 04-DEC-1998; 98US-0206059.
XX
XX (ENTR-) ENTREMED INC.
XX
XX Macdonald NJ, Sim KL;
XX
XX WPI; 2000-412290/35.
XX
XX New angiogenesis-inhibiting protein receptors, useful in methods for
PT treating diseases and processes that are mediated by angiogenesis, such
PT as solid tumours, psoriasis, scleroderma and myocardial angiogenesis -

```

XX Disclosure; Page 86; 100pp; English.

XX PS

XX This invention relates to angiogenesis-inhibiting protein receptors, and the DNA sequences encoding them. Angiogenesis is the generation of new blood vessels into a tissue, and normally occurs in wound healing, fetal and embryonal development, and the formation of the corpus luteum, endometrium and placenta. Angiostatin is a protein (see AAB16450 and AAA68202) involved in angiogenesis, and has an amino acid sequence similar to that of a plasminogen fragment (see murine plasminogen AAB16490). Angiostatin has the ability to inhibit angiogenesis.

XX Endostatin is also an angiogenesis inhibiting protein (see AAB16451 and AAA68203). Sequences AAB68242 and AAB16522 represent coding and protein sequences of human laminin. Laminin is an angiostatin binding protein, and some of the peptides of the invention share homology with regions of laminin. Peptides AAB16452-B16521 (excluding AAB16490) are the angiogenesis-inhibiting protein receptor fragments of the invention. The peptides bind either angiostatin or endostatin and can be used in methods for treating diseases and processes that are mediated by angiogenesis, such as solid tumours, psoriasis, scleroderma, myocardial angiogenesis, Crohn's disease, cerebral collaterals, arteriovenous malformations, rubeosis, diabetic retinopathy, arthritis, wound healing, peptic ulcers, Helicobacter related diseases, fractures, placentalion and cat scratch fever. They are useful for the detection and prognosis of cancer. DNA sequences A628204-A628241 encode the peptides of the invention.

XX PA

XX Sequence 552 BP; 83 A; 196 C; 179 G; 94 T; 0 other;

XX PI

XX Query Match 75.3%; Score 418; DB 21; Length 552;

XX DR Best Local Similarity 85.3%; Pred. No. 3.4e-62;

XX DR P-PSDB; AAB49379.

XX Mismatches 0; Indels 0; Gaps 0;

XX Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 1 CACACCCACAGGACTTCCAGCTGGTGTGCTGACCTGGTGGCCCTGAACAGCCCGACGCCG 60

DB 1 CACAGCCACCGGACTTCCAGCCGGTGTCTCCACCTGGTGGCTCAACAGCCCTGTCA 60

QY 61 GCGCGATCGAGGATCCCGGGAGCGGACTTCCAGTGTCTTCAGAGCGCGCGCGCG 120

DB 61 GCGCGATCGAGGATCCCGGGAGCGGACTTCCAGTGTCTTCAGAGCGCGCGCGCG 120

QY 121 GGGCTGGCGGACCTTCCGGGCGTTCCTGCTGGCGCTTCCAGTGTCTTCAGAGCATC 180

DB 121 GGGCTGGCGGACCTTCCGGGCGTTCCTGCTGGCGCTTCCAGTGTCTTCAGAGCATC 180

QY 181 GTGGCGCGCGGACCTTCCGGGCGTTCCTGCTGGCGCTTCCAGTGTCTTCAGAGCATC 240

DB 181 GTGGCGCGCGGACCTTCCGGGCGTTCCTGCTGGCGCTTCCAGTGTCTTCAGAGCATC 240

QY 241 CCGAGTGGAGGCTTATTCCTGGGCTCCGAGGCGGAGCTGAAGCCCGGGCGCGCATC 300

DB 241 CCGAGTGGAGGCTTATTCCTGGGCTCCGAGGCGGAGCTGAAGCCCGGGCGCGCATC 300

QY 301 TTCTCTTTCAGGCGAGAGTGTCTGAGCAGCCCGCTGCTGCGGCGGAGCGGTGG 360

DB 301 TTCTCTTTCAGGCGAGAGTGTCTGAGCAGCCCGCTGCTGCGGCGGAGCGGTGG 360

QY 361 CAGCGCTCCGACCCAGCGGGCGCGCTTCCAGGCGGAGCTACTGAGAGCGTGGGACG 420

DB 361 CAGCGCTCCGACCCAGCGGGCGCGCTTCCAGGCGGAGCTACTGAGAGCGTGGGACG 420

QY 421 GAGGCGCGGCGGACCGCGGCGAGGCTGTCTGCTGGCGGCGAGGCTGTGGAGCAG 480

DB 421 GAGGCTCCCTCGGCGGCGGCGGCGGCTTCTGCTGGCGGCGAGGCTGTGGAGCAG 480

QY 481 GAGGCGCGGCGGCGGCGGCGGCGGCTTCTGCTGGCGGCGAGGCTGTGGAGCAG 540

DB 481 GAGGCGCGGCGGCGGCGGCGGCGGCTTCTGCTGGCGGCGAGGCTGTGGAGCAG 540

QY 541 TCCTTC 546

DB 541 GCCTCC 546

XX PS

XX This invention provides endostatin peptides which can be used in the modulation of angiogenesis. This is useful in the treatment of cancers, inflammation, rheumatoid arthritis, chronic articular rheumatism, psoriasis, disorders associated with inopportune invasion of vessels such as diabetic retinopathy, neovascular glaucoma, retinopathy of prematurity, macular degeneration, corneal graft rejection, retrolental fibroplasia, rubeosis, capillary proliferation in atherosclerotic plaques and osteoporosis. Other angiogenesis-dependent diseases include Osler-Webber syndrome, myocardial angiogenesis, plaque neovascularisation, telangiectasia, haemophiliac joints and wound granulation. In addition, the peptides can be used as birth control agents.

XX PA

XX Sequence 552 BP; 83 A; 196 C; 179 G; 94 T; 0 other;

XX PI

XX Query Match 75.3%; Score 418; DB 22; Length 552;

XX DR Best Local Similarity 85.3%; Pred. No. 3.4e-62;

XX DR P-PSDB; AAB49379.

XX Mismatches 0; Indels 0; Gaps 0;

XX Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 1 CACACCCACAGGACTTCCAGCTGGTGTGCTGACCTGGTGGCCCTGAACAGCCCGACGCCG 60

DB 1 CACAGCCACCGGACTTCCAGCCGGTGTCTCCACCTGGTGGCTCAACAGCCCTGTCA 60

QY 61 GCGCGCATCGAGGATCCCGGGAGCGGACTTCCAGTGTCTTCAGAGCGCGCGCGCG 120

DB 61 GCGCGCATCGAGGATCCCGGGAGCGGACTTCCAGTGTCTTCAGAGCGCGCGCGCG 120

QY 121 GGGCTGGCGGACCTTCCGGGCGTTCCTGCTGGCGCTTCCAGTGTCTTCAGAGCATC 180

DB 121 GGGCTGGCGGACCTTCCGGGCGTTCCTGCTGGCGCTTCCAGTGTCTTCAGAGCATC 180

QY 181 GTGGCGCGCGGACCTTCCGGGCGTTCCTGCTGGCGCTTCCAGTGTCTTCAGAGCATC 240

DB 181 GTGGCGCGCGGACCTTCCGGGCGTTCCTGCTGGCGCTTCCAGTGTCTTCAGAGCATC 240

QY 241 CCGAGTGGAGGCTTATTCCTGGGCTCCGAGGCGGAGCTGAAGCCCGGGCGCGCATC 300

DB 241 CCGAGTGGAGGCTTATTCCTGGGCTCCGAGGCGGAGCTGAAGCCCGGGCGCGCATC 300

QY 301 TTCTCTTTCAGGCGAGAGTGTCTGAGCAGCCCGCTGCTGCGGCGGAGCGGTGG 360

DB 301 TTCTCTTTCAGGCGAGAGTGTCTGAGCAGCCCGCTGCTGCGGCGGAGCGGTGG 360

QY 361 CAGCGCTCCGACCCAGCGGGCGCGCTTCCAGGCGGAGCTACTGAGAGCGTGGGACG 420

DB 361 CAGCGCTCCGACCCAGCGGGCGCGCTTCCAGGCGGAGCTACTGAGAGCGTGGGACG 420

QY 421 GAGGCGCGGCGGACCGCGGCGAGGCTGTCTGCTGGCGGCGAGGCTGTGGAGCAG 480

DB 421 GAGGCTCCCTCGGCGGCGGCGGCGGCTTCTGCTGGCGGCGAGGCTGTGGAGCAG 480

QY 481 GAGGCGCGGCGGCGGCGGCGGCGGCTTCTGCTGGCGGCGAGGCTGTGGAGCAG 540

DB 481 GAGGCGCGGCGGCGGCGGCGGCGGCTTCTGCTGGCGGCGAGGCTGTGGAGCAG 540

QY 541 TCCTTC 546

DB 541 GCCTCC 546

XX PS

XX This invention provides endostatin peptides which can be used in the modulation of angiogenesis. This is useful in the treatment of cancers, inflammation, rheumatoid arthritis, chronic articular rheumatism, psoriasis, disorders associated with inopportune invasion of vessels such as diabetic retinopathy, neovascular glaucoma, retinopathy of prematurity, macular degeneration, corneal graft rejection, retrolental fibroplasia, rubeosis, capillary proliferation in atherosclerotic plaques and osteoporosis. Other angiogenesis-dependent diseases include Osler-Webber syndrome, myocardial angiogenesis, plaque neovascularisation, telangiectasia, haemophiliac joints and wound granulation. In addition, the peptides can be used as birth control agents.

XX PA

XX Sequence 552 BP; 83 A; 196 C; 179 G; 94 T; 0 other;

XX PI

XX Query Match 75.3%; Score 418; DB 22; Length 552;

XX DR Best Local Similarity 85.3%; Pred. No. 3.4e-62;

XX DR P-PSDB; AAB49379.

XX Mismatches 0; Indels 0; Gaps 0;

XX Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 1 CACACCCACAGGACTTCCAGCTGGTGTGCTGACCTGGTGGCCCTGAACAGCCCGACGCCG 60

DB 1 CACAGCCACCGGACTTCCAGCCGGTGTCTCCACCTGGTGGCTCAACAGCCCTGTCA 60

QY 61 GCGCGATCGAGGATCCCGGGAGCGGACTTCCAGTGTCTTCAGAGCGCGCGCGCG 120

DB 61 GCGCGATCGAGGATCCCGGGAGCGGACTTCCAGTGTCTTCAGAGCGCGCGCGCG 120

QY 121 GGGCTGGCGGACCTTCCGGGCGTTCCTGCTGGCGCTTCCAGTGTCTTCAGAGCATC 180

DB 121 GGGCTGGCGGACCTTCCGGGCGTTCCTGCTGGCGCTTCCAGTGTCTTCAGAGCATC 180

QY 181 GTGGCGCGCGGACCTTCCGGGCGTTCCTGCTGGCGCTTCCAGTGTCTTCAGAGCATC 240

DB 181 GTGGCGCGCGGACCTTCCGGGCGTTCCTGCTGGCGCTTCCAGTGTCTTCAGAGCATC 240

QY 241 CCGAGTGGAGGCTTATTCCTGGGCTCCGAGGCGGAGCTGAAGCCCGGGCGCGCATC 300

DB 241 CCGAGTGGAGGCTTATTCCTGGGCTCCGAGGCGGAGCTGAAGCCCGGGCGCGCATC 300

QY 301 TTCTCTTTCAGGCGAGAGTGTCTGAGCAGCCCGCTGCTGCGGCGGAGCGGTGG 360

DB 301 TTCTCTTTCAGGCGAGAGTGTCTGAGCAGCCCGCTGCTGCGGCGGAGCGGTGG 360

QY 361 CAGCGCTCCGACCCAGCGGGCGCGCTTCCAGGCGGAGCTACTGAGAGCGTGGGACG 420

DB 361 CAGCGCTCCGACCCAGCGGGCGCGCTTCCAGGCGGAGCTACTGAGAGCGTGGGACG 420

QY 421 GAGGCGCGGCGGACCGCGGCGAGGCTGTCTGCTGGCGGCGAGGCTGTGGAGCAG 480

DB 421 GAGGCTCCCTCGGCGGCGGCGGCGGCTTCTGCTGGCGGCGAGGCTGTGGAGCAG 480

QY 481 GAGGCGCGGCGGCGGCGGCGGCGGCTTCTGCTGGCGGCGAGGCTGTGGAGCAG 540

DB 481 GAGGCGCGGCGGCGGCGGCGGCGGCTTCTGCTGGCGGCGAGGCTGTGGAGCAG 540

QY 541 TCCTTC 546

DB 541 GCCTCC 546

XX PS

XX This invention provides endostatin peptides which can be used in the modulation of angiogenesis. This is useful in the treatment of cancers, inflammation, rheumatoid arthritis, chronic articular rheumatism, psoriasis, disorders associated with inopportune invasion of vessels such as diabetic ret

```
Db 241 CCAGCTGGGAGGCTGTTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGCAGCATC 300
QY 301 TTCTCTTTGACGGCAGAGATGCTTCGACGACCCCGCCCTGGCCCGGAGAGCGTGTGG 360
Db 301 TTCTCTTTGACGGCAGAGAGCTCTGAGCGACCCACCTGGCCCGGAGAGCGTGTGG 360
QY 361 CACGGCTCCGACCCCGAGCGGGCCGCTCAGCGACAGCTACTGCGAGAGCTGGCGGACG 420
Db 361 CATGGCTCGGACCCCAACGGGCGAGGCTCAGCGAGCTACTGTGAGAGCTGGCGGACG 420
QY 421 GAGGCCCCCGCGGCCACCGGGCAGGCGTCTGCTGCTGGCGGCGAGGCTGCTGGAGCAG 480
Db 421 GAGGCTCCCTCGGCCACCGGCCAGGCTCTGCTGCTGGGGCGAGGCTCTCTGGGGCAG 480
QY 481 GAGGCGCGAGCTGCGGCCACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 AGTGGCGGAGCTGCCATCAGGCTACATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 TCCTTC 546
Db 541 GCCTCC 546
```

RESULT 15

ABK50685
ID ABK50685 standard; cDNA; 552 BP.

AC ABK50685;

XX 13-AUG-2002 (first entry)

DT cDNA encoding human endostatin.

DE
XX Human; angiogenesis; PITSRE protein kinase; cancer; arthritis;
KW macular degeneration; diabetic retinopathy; angiogenic-related disease;
KW haemangioma; blood borne tumour; leukaemia; neovascularisation;
KW coronary collateral; cerebral collateral; neovascular glaucoma;
KW corneal disease; wound healing; Helicobacter related disease; fracture;
KW keloid; ovulation; menstruation; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FH 1..552

FT /*tag= a

FT /partial

FT /product= "Endostatin"

FT /note= "This sequence lacks a start codon"

XX WO200230982-A2.

PN 18-APR-2002.

XX 15-OCT-2001; 2001WO-US32437.

XX 13-OCT-2000; 2000US-240127P.

XX (ENTR-) ENTREMED INC.

XX Sim KL, Liang H;

XX WPI; 2002-435440/46.

DR P-PSDB; AAU97132.

XX Regulating angiogenesis for treating scleroderma, leukaemia, keloids by
PT administering a protein that is homologous to PITSRE protein kinase
PT and an angiogenic factor or a protein kinase and its active fragments -

XX Disclosure; Fig 2B; 45pp; English.

XX The present invention relates to methods and compositions of inhibiting
CC angiogenesis. The method comprises administering to a human or animal

CC a composition comprising a protein that is homologous to PITSRE
CC protein kinases (PK) and an angiogenic factor. The method is useful
CC for regulating angiogenesis related to cancer, arthritis, macular
CC degeneration, and diabetic retinopathy. The compositions are useful
CC for inhibiting angiogenic-related diseases. The method and
CC compositions are useful in treating diseases and processes that are
CC mediated by angiogenesis including haemangioma, solid tumours, blood
CC borne tumours, leukaemia, metastasis, telangiectasia, psoriasis, scleroderma, pyogenic granuloma, myocardial angiogenesis, Crohn's
CC disease, plaque neovascularisation, coronary collaterals, cerebral
CC collaterals, arteriovenous malformations, ischaemic limb angiogenesis,
CC corneal diseases, rubeosis, neovascular glaucoma, diabetic retinopathy,
CC retrolental fibroplasia, arthritis, diabetic neovascularisation,
CC macular degeneration, wound healing, peptic ulcer, Helicobacter related
CC diseases, fractures, keloids, vasculogenesis, haematopoiesis, ovulation,
CC menstruation, placentaion, and cat scratch fever. The method of
CC the invention provides a therapy for cancer that has minimal side
CC effects. The present sequence encodes human endostatin which is
CC used to generate angiogenesis-inhibiting peptides.

XX
SQ Sequence 552 BP; 83 A; 196 C; 179 G; 94 T; 0 other;

Query Match 75.3%; Score 418; DB 24; Length 552;

Best Local Similarity 85.3%; Pred. No. 3.4e-62;

Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 1 CACACCCACGAGCTTCCAGCTGGTGTGACCTGGTGGCCCTGAACAGCCGCGAGCG 60

Db 1 CACAGCCACCGGAGCTTCCAGCCGGTGTCTCCACCTGGTTCACAGCCCCCTGTCA 60

QY 61 GCGGCGATGCGAGGCGATCCGGGGAGCGGACTTCCAGTGTCTTCAGCAGCGCGCGCG 120

Db 61 GCGGCGATGCGGGGCGATCCGGGGGCGGACTTCCAGTGTCTTCAGCAGCGCGCGCG 120

QY 121 GGGCTGGCGCGCACCTTCCGGGCGCTTCTGTGCTGCGCGCTGCGAGACCTTACAGCATC 180

Db 121 GGGCTGGCGCGCACCTTCCGGGCGCTTCTGTGCTGCGCGCTGCGAGACCTGTCAGCATC 180

QY 181 GTGCGCGCGCGCGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240

Db 181 GTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240

QY 241 CCCAGCTGGGAGGCGCTTATTTCTGGGCTCCGAGGCGCAGCTGAAGCCCGGGCGCGCATC 300

Db 241 CCCAGCTGGGAGGCGCTTATTTCTGGGCTCCGAGGCGCAGCTGAAGCCCGGGCGCGCATC 300

QY 301 TTCTCTTTGCGACGGCAGAGATGCTCTGCGAGCACCCTGCGCGCGGCGGAGAGCGTGTGG 360

Db 301 TTCTCTTTGCGACGCAAGGAGCGTCTCTGAGGCGACCCACCTGGCGCCCGAGAGCGTGTGG 360

QY 361 CAGGCTCCGACCCCGAGCGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420

Db 361 CATGGCTCGGACCCCAACGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420

QY 421 GAGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480

Db 421 GAGGCTCCCTCGGCCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480

QY 481 GAGGCGCGGAGCTGCGCGCCACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

Db 481 AGTGGCGGAGCTGGCATCACGCTACATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 540

QY 541 TCCTTC 546

Db 541 GCCTCC 546

Search completed: August 17, 2003, 13:19:26

Job time : 214.739 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 13:57:19 ; Search time 56.5426 Seconds
(without alignments)
4332.442 Million cell updates/sec

Title: US-09-938-391-3
Perfect score: 555
Sequence: 1 cacaccaccaggactcca.....tgacctcttctccaagtag 555

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/2/ina/5A_COMB.seq:
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:
3: /cgn2_6/prodata/2/ina/6A_COMB.seq:
4: /cgn2_6/prodata/2/ina/6B_COMB.seq:
5: /cgn2_6/prodata/2/ina/PCFUS_COMB.seq:
6: /cgn2_6/prodata/2/ina/backfiles1.seq:

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	418	75.3	546	4	US-09-315-689-4
2	418	75.3	552	3	US-09-206-059-30
3	418	75.3	3394	1	US-08-159-784-4
4	410.8	74.0	534	4	US-09-315-689-6
5	375.8	67.7	558	3	US-09-449-293-3
6	375.8	67.7	558	4	US-09-775-325-3
7	375.8	67.7	565	3	US-08-985-526-37
8	374.2	67.4	4031	1	US-08-159-784-1
9	373.4	67.3	573	4	US-09-561-500-12
10	373.4	67.3	573	4	US-09-561-108-12
11	373.4	67.3	573	4	US-09-561-526-12
12	373.4	67.3	573	4	US-09-561-499-12
13	60.8	11.0	1116	4	US-09-252-991A-2102
C 14	60.8	11.0	1224	4	US-09-252-991A-1793
C 15	59.4	10.7	1260	4	US-09-252-991A-1861
C 16	59.4	10.7	1722	4	US-09-252-991A-2173
17	57.4	10.3	741	4	US-09-252-991A-6774
C 18	57.4	10.3	1557	4	US-09-252-991A-6920
C 19	57.4	10.3	1590	4	US-09-252-991A-6824
C 20	55.8	10.1	1820	1	US-08-173-508-7
C 21	55.8	10.1	1821	2	US-08-265-310-7
C 22	55.8	10.1	1821	3	US-08-961-742-7
C 23	53.8	9.7	1050	4	US-09-252-991A-10973
C 24	53.8	9.7	2547	4	US-09-252-991A-11271
C 25	53.8	9.7	2964	4	US-09-252-991A-10912
26	53.6	9.7	1140	4	US-09-522-714-13
27	53	9.5	711	4	US-09-252-991A-11047

28	53	9.5	15872	3	US-09-105-537-1	Sequence 1, Appli
29	52.4	9.4	30001	1	US-08-125-468-1	Sequence 1, Appli
30	52.4	9.4	30001	2	US-08-474-933-1	Sequence 1, Appli
31	52	9.4	20235	1	US-07-642-734C-3	Sequence 3, Appli
32	52	9.4	20235	3	US-08-439-009A-3	Sequence 3, Appli
C 33	51.8	9.3	1260	4	US-09-252-991A-8957	Sequence 8957, Ap
34	51.8	9.3	1314	4	US-09-252-991A-9048	Sequence 9048, Ap
35	51.8	9.3	2133	4	US-09-252-991A-9129	Sequence 9129, Ap
C 36	49.4	8.9	71989	4	US-09-443-501A-2	Sequence 2, Appli
C 37	49.4	8.9	411529	3	US-09-103-840A-1	Sequence 1, Appli
38	49	8.8	2143	2	US-08-656-177A-1	Sequence 1, Appli
39	49	8.8	2143	3	US-09-256-797-1	Sequence 1, Appli
40	48.4	8.7	23673	4	US-09-773-816-1	Sequence 1, Appli
C 41	48.2	8.7	447	4	US-09-252-991A-1765	Sequence 1765, Ap
C 42	48.2	8.7	1071	4	US-09-252-991A-15521	Sequence 15521, A
C 43	48.2	8.7	1371	4	US-09-252-991A-15313	Sequence 15313, A
C 44	48.2	8.7	1701	4	US-09-252-991A-1960	Sequence 1960, Ap
C 45	48.2	8.7	2169	4	US-09-252-991A-1910	Sequence 1910, Ap

ALIGNMENTS

RESULT 1
US-09-315-689-4
; Sequence 4, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Folkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-315-689-4

Query Match 75.3%; Score 418; DB 4; Length 546;
Best Local Similarity 85.3%; Pred. No. 7.6e-74;
Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY	1	CACACCCACGAGACTTCAGCTGGTGTGACCTGGTGGCCCTGAAAGCCGCGAGCGG	60
DB	1	CACAGCCACCGGACTTCCAGCCGGTGTCTCCACCTGGTTGGCTCAACAGCCCCCTGTCA	60
QY	61	GGCGCATCGAGGATCCGGGAGCGGACTTCCAGTGTCTCCAGGCGCGCGCGG	120
DB	61	GGCGCATCGGGGATCCGGGGCGGACTTCCAGTGTCTCCAGGCGCGCGCGG	120
QY	121	GGGCTGGCGGACCTTCCGGGCTTCTGTGCTCGCGCTCGAGGACCTCTACAGCATC	180
DB	121	GGGCTGGCGGACCTTCCGGGCTTCTGTGCTCGCGCTCGAGGACCTGTACAGCATC	180
QY	181	GTGCGCGCGCGACCGACCGGGGTGCGGCTGCTCAACTCAGGAGCAGAGTGTCTTTC	240
DB	181	GTGCGCGGTGCGGACCGCGGCGGCTGCGGCTGCTCAACTCAGGAGCAGAGTGTCTTTC	240
QY	241	CCAGCTGGAGGCTTATTTCTCGGCTCCGAGGCGGCTGAGGCGGCGGCGGCGGCGG	300
DB	241	CCAGCTGGAGGCTTATTTCTCGGCTCCGAGGCGGCTGAGGCGGCGGCGGCGGCGG	300
QY	301	TTCTCTTTTCGAGCGAGAGATGCTCTGAGCAGCCCGCTGGCCCGGAGAGCGGTGTGG	360
DB	301	TTCTCTTTTCGAGCGAGAGAGATGCTCTGAGCAGCCCGCTGGCCCGGAGAGCGGTGTGG	360
QY	361	CACGCTCCGACCCCGAGCGGCGGCGGCTGACCGAGCTACTGCGAGACGTGGCGGAGCG	420
DB	361	CATGCTCGGACCCCAACGGGCGGAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGAGCG	420

QY 541 TCCTTCTCCAAGTAG 555
|||
Db 544 TCTTCTCCAATAG 558

RESULT 6

US-09-775-325-3
; Sequence 3, Application US/09775325
; Patent No. 6500449
; GENERAL INFORMATION:
; APPLICANT: Adibol, Marc
; APPLICANT: Uteza, Yves
; APPLICANT: Menasche, Maurice
; APPLICANT: Bossard, Carine
; APPLICANT: Van Den Bergh, Loic
; APPLICANT: Bonnel, Sebastien
; APPLICANT: Prats, Hervé
; APPLICANT: Honiger, Jiri
; APPLICANT: Neuner-Jehle, Martin
; TITLE OF INVENTION: INTRAOCULAR TRANSPLANTATION OF ENCAPSULATED CELLS
; FILE REFERENCE: 8076.202USD1
; CURRENT APPLICATION NUMBER: US/09/775,325
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: US 09/449,293
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Rattus rattus
US-09-775-325-3

Query Match 67.7%; Score 375.8; DB 4; Length 558;

Best Local Similarity 79.8%; Pred. No. 1.4e-65;
Matches 443; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 1 CACACCCACGAGCTTCCAGCTGGTGTGCTGACCTGGTGGCCCTGAACAGCCGCGAGCG 60
Db 4 CATACTCATGAGACTTTCAGCCAGTGTCTCCACTGGTGGCACTGAACACCCCTGTCT 63
QY 61 GCGCGCATGCGAGGCATCCGCGGAGCGGAGCTTCACGTGCTTCCAGCAGCGCGCGCGG 120
Db 64 GGAGGATCGGTGGTATCCGTGAGCAGATTTCCAGTGTCTTCCAGCAGCCGAGCGG 123
QY 121 GGGCTGGCGGACCTTCCGCGGCTTCTGTGCTGCTGCGGCTGAGACCTCTACAGCATC 180
Db 124 GGGCTGTGCGGCACCTTCCGCGGCTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 183
QY 181 GTGCGCGCGCGCACCTTCCGCGGCTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 184 GTGCGCGCGTGTGCTGCGGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
QY 241 CCAGCTGGGAGGCTTATTTCTCGGCTCCGAGGCGCAGCTGAAGCCCGGGCCCGCATC 300
Db 244 CCAGCTGGGAGCTCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 303
QY 301 TTCTCTTTGACGGCAGAGATGCTTCTGAGCAGACCCCGCTGCGCCCGGAGAGCGTGTG 360
Db 304 TTTTCTTTTGAAGCAGAGATGCTTCTGAGCAGACCCAGCTGCGCGCAGAGAGCGTATG 363
QY 361 CAGGCTCCGACCCGAGCGGCGCGCTGACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 364 CAGGCTCCGACCCGAGCGGCGCGCTGACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
QY 421 GAGGCGCGCGCGCACCCGCGGAGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 424 GAACTACTGGGCTACAGTCTGAGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483
QY 481 GAGGCGCGAGCTGCGCGCACCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 484 AAAGCTGCGAGCTGCCACACAGCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543

QY 541 TCCTTCTCCAAGTAG 555
|||
Db 544 TCTTCTCCAATAG 558

RESULT 7

US-08-985-526-37
; Sequence 37, Application US/08985526
; Patent No. 6080728
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA
; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Connolly, Bove, Lodge, & Hutz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985.526
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorow Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 565 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-985-526-37

Query Match 67.7%; Score 375.8; DB 3; Length 565;

Best Local Similarity 79.8%; Pred. No. 1.4e-65;
Matches 443; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 1 CACACCCACGAGCTTCCAGTGTGCTGCACTGGTGGCCCTGAACAGCCGCGAGCG 60
Db 11 CATACTCATGAGACTTTCAGCCAGTGTCTCCACTGGTGGCACTGAACACCCCTGTCT 70
QY 61 GCGCGCATGCGAGGCATCCGCGGAGCGGACTTCCAGTGTCTTCCAGCAGCGCGCGCG 120
Db 71 GGAGGATCGGTGGTATCCGTGAGCAGATTTCCAGTGTCTTCCAGCAGCCGCGCGTG 130
QY 121 GGGCTGGCGCGCACCTTCCGGGCTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 131 GGGCTGTGCGGCACCTTCCGGGCTTCTGTGCTCTTAGGCTGAGGATCTCTATAGCATC 190
QY 181 GTGCGCGCGCGCACCGCACCGGGGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 191 GTGCGCGCGTGTGACCGCGGGGTCTGTGCGCATCTGCTCAACCTGAAGAGCAGGTGCT 250
QY 241 CCCAGCTGGAGGCTTATTTCTCGGCTCCGAGGCGCAGCTGAAGCCCGGGCCCGCATC 300
Db 251 CCCAGCTGGAGCTCCCTGTTTCTGGCTCCAGGGTCAACTGAACCCCGGGCCCGCATC 310
QY 301 TTCTCTTTTCGAGCGCAGAGATGCTCTGAGCAGCACCCCGCTGCGCCCGGAGAGCGTGTG 360

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 14:02:22 ; Search time 208.125 Seconds
(without alignments)
5967.841 Million cell updates/sec

Title: US-09-938-391-3
Perfect score: 555
Sequence: 1 cacaccaccagacttcca.....tgacctcttctccaagtag 555

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 1504479 seqs, 1118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
12: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	550.4	99.2	552	12	US-10-292-418-34
2	550.4	99.2	552	14	US-10-131-241-50
3	418	75.3	546	14	US-10-042-347-4
4	418	75.3	549	12	US-10-292-418-3
5	418	75.3	549	14	US-10-131-241-53
6	418	75.3	552	9	US-09-873-676-30
7	418	75.3	3394	10	US-09-880-107-2178
8	418	75.3	4551	14	US-10-060-036-144
9	417	75.1	551	13	US-10-080-797-2
10	415.6	74.9	632	14	US-10-131-241-51
11	410.8	74.0	534	14	US-10-042-347-6
12	410.8	74.0	537	14	US-10-131-241-59
13	403	72.6	540	14	US-10-131-241-48
14	375.8	67.7	558	9	US-09-775-174-3
15	375.8	67.7	558	9	US-09-775-325-3
16	375.8	67.7	565	13	US-10-036-869-37

17	374.2	67.4	624	13	US-10-080-797-4	Sequence 4, Appli
18	373.4	67.3	573	10	US-09-998-831-12	Sequence 12, Appl
19	371.8	67.0	552	12	US-10-292-418-17	Sequence 17, Appl
20	187.8	33.8	574	14	US-10-060-036-64	Sequence 64, Appl
21	171.2	30.8	5204	12	US-09-971-392-159	Sequence 159, App
22	95.4	17.2	900	14	US-10-131-241-45	Sequence 45, Appl
23	74.4	13.4	434	14	US-10-060-036-4	Sequence 4, Appli
24	58.4	10.5	1728	14	US-10-156-761-4493	Sequence 4493, Ap
25	58.4	10.5	9025608	14	US-10-156-761-1	Sequence 1, Appli
26	57	10.3	1548	14	US-10-156-761-3078	Sequence 3078, Ap
27	57	10.3	9025608	14	US-10-156-761-1	Sequence 1, Appli
28	56.2	10.1	2079	14	US-10-076-816-3	Sequence 3, Appli
29	56.2	10.1	2079	14	US-10-077-381-3	Sequence 3, Appli
30	56	10.1	1785	14	US-10-156-761-7479	Sequence 7479, Ap
31	55.8	10.1	1821	14	US-10-156-761-3274	Sequence 3274, Ap
32	55.6	10.0	1704	14	US-10-156-761-3928	Sequence 3928, Ap
33	54.6	9.8	1230	14	US-10-156-761-5404	Sequence 5404, Ap
34	54	9.7	1359	12	US-10-148-907A-14	Sequence 14, Appl
35	53.8	9.7	1428	14	US-10-156-761-4949	Sequence 4949, Ap
36	53.6	9.7	1140	14	US-10-304-928-13	Sequence 13, Appl
37	53.4	9.6	1029	11	US-09-938-901-5	Sequence 5, Appli
38	53.4	9.6	3483	14	US-10-156-761-6532	Sequence 6532, Ap
39	53.4	9.6	88421	10	US-09-976-059-1	Sequence 1, Appli
40	53.2	9.6	1941	14	US-10-156-761-3418	Sequence 3418, Ap
41	53	9.5	15872	10	US-09-861-289-1	Sequence 1, Appli
42	53	9.5	15872	11	US-09-860-846-1	Sequence 1, Appli
43	53	9.5	15872	11	US-09-988-384B-1	Sequence 1, Appli
44	53	9.5	15872	11	US-09-836-821-1	Sequence 1, Appli
45	52.8	9.5	924	14	US-10-156-761-1189	Sequence 1189, Ap

ALIGNMENTS

RESULT 1

US-10-292-418-34
; Sequence 34, Application US/10292418
; Publication No. US20030139365A1
; GENERAL INFORMATION:
; APPLICANT: Lo, Kin-Ming
; APPLICANT: Li, Yue
; APPLICANT: Gillies, Stephen D
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
; TITLE OF INVENTION: Immunofusins
; FILE REFERENCE: LEX-006C1
; CURRENT APPLICATION NUMBER: US/10/292,418
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/383,315
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: US 60/097,883
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(552)
; OTHER INFORMATION: Endostatin
US-10-292-418-34

Query Match 99.2%; Score 550.4; DB 12; Length 552;
Best Local Similarity 99.8%; Pred. No. 7.7e-117;
Matches 551; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CACACCCACGAGACTTCCAGCTGGTGTGACCTGGTGGCTGAAACAGCCGCGCG 60
Db 1 CACACCCACGAGACTTCCAGCTGGTGTGACCTGGTGGCTGAAACAGCCGCGCG 60
QY 61 GCGCGCATCGAGGATCCGGGAGCGGACTTCCAGTGTTCAGCAGCGCGCGCGCG 120

Db	61	GGCGGATCGAGGCATCGGGGACGGACCTTCAGTGTCTTCAGACGAGCGCGCGCGG	120
Qy	121	GGGCTGGCGGCACCTTCGGGCTTCCTGTCGTCGGGCTGCAGGACCTCTACAGCATC	180
Db	121	GGGCTGGCGGCACCTTCGGGCTTCCTGTCGTCGGGCTGCAGGACCTCTACAGCATC	180
Qy	181	GTGGCGCGCGCGACGGACCGGGTGCCTGTCGTCACCTCAGGAGCAGGTGCTCTTC	240
Db	181	GTGGCGCGCGCGACCGACCGGGTGCCTGTCGTCACCTCAGGAGCAGGTGCTCTTC	240
Qy	241	CCAGCTGGGAGGCGCTTATCTCGGGCTCCAGGGCCAGCTGAAGCCCGGGGCCGCATC	300
Db	241	CCAGCTGGGAGGCGCTTATCTCGGGCTCCAGGGCCAGCTGAAGCCCGGGGCCGCATC	300
Qy	301	TTCTCTTTCAGCGCAGAGATGTCTCAGACACCCCGCTGCGCCCGGAAGAGCGTGTGG	360
Db	301	TTCTCTTTCAGCGCAGAGATGTCTCAGACACCCCGCTGCGCCCGGAAGAGCGTGTGG	360
Qy	361	CACGGCTCCGACCCAGCGGGCGCGCTGACGCACAGCTACTGCGAGACGTGGCGGACG	420
Db	361	CACGGCTCCGACCCAGCGGGCGCGCTGACGCACAGCTACTGCGAGACGTGGCGGACG	420
Qy	421	GAGGCGCGCGCGCACCCGGGAGCGCTGCTGCTGTGCGGCGAAGGCTGTGGAGCAG	480
Db	421	GAGGCGCGCGCGCACCCGGGAGCGCTGCTGCTGTGCGGCGAAGGCTGTGGAGCAG	480
Qy	481	GAGGCGCGAGCTCGCGCCACCGCTTGTGTGTGCTTGCATCGAGACAGCGTCATGCC	540
Db	481	GAGGCGCGAGCTCGCGCCACCGCTTGTGTGTGCTTGCATCGAGACAGCGTCATGCC	540
Qy	541	TCCTTCTCCAAG	552
Db	541	TCCTTCTCCAAG	552

RESULT 2
 US-10-131-241-50
 ; Sequence 50, Application US/10131241
 ; Publication No. US20030012792A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Holaday, John W.
 ; APPLICANT: Fortier, Anne H.
 ; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
 ; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
 ; FILE REFERENCE: 052113-0344 43170-271565
 ; CURRENT APPLICATION NUMBER: US/10/131,241
 ; CURRENT FILING DATE: 2002-07-22
 ; PRIOR APPLICATION NUMBER: US 09/413,049
 ; PRIOR FILING DATE: 1999-10-06
 ; PRIOR APPLICATION NUMBER: US 09/316,802
 ; PRIOR FILING DATE: 1999-05-21
 ; PRIOR APPLICATION NUMBER: US 60/086,586
 ; PRIOR FILING DATE: 1998-05-22
 ; NUMBER OF SEQ ID NOS: 65
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 50
 ; LENGTH: 552
 ; TYPE: DNA
 ; ORGANISM: Canine sp.
 US-10-131-241-50

Qy	121	GGGCTGGCGGCACCTTCGCGGCCTTCTCTGTCGCGGCTGCAGGACCTCTACAGCATC	180
Db	121	GGGCTGGCGGCACCTTCGCGGCCTTCTCTGTCGCGGCTGCAGGACCTCTACAGCATC	180
Qy	181	GTGGCGCGCGCGACGCACACCGGGGTGCGCGTCTCAACCTCAGGAGCAGGTGCTCTTC	240
Db	181	GTGGCGCGCGCGACCGCACCGGGGTGCCCGTCTCAACCTCAGGAGCAGGTGCTCTTC	240
Qy	241	CCCAGCTGGGAGGCGCTTATTCTCGGGCTCCGAGGCGCAGCTGAAGCCCGGGGCCCGCATC	300
Db	241	CCCAGCTGGGAGGCGCTTATTCTCGGGCTCCGAGGCGCAGCTGAAGCCCGGGGCCCGCATC	300
Qy	301	TTCTCTTTTCAGCGCAGAGATGCTCGACACACCGCCTCGGCCCGCGAAGAGCGTGTGG	360
Db	301	TTCTCTTTTCAGCGCAGAGATGCTCGACACACCGCCTCGGCCCGCGAAGAGCGTGTGG	360
Qy	361	CACGGCTCCGACCCCGAGCGGCGCGGCTGACCGACAGCTACTGCGACACGTGGCGGAGC	420
Db	361	CACGGCTCCGACCCCGAGCGGCGCGGCTGACCGACAGCTACTGCGACACGTGGCGGAGC	420
Qy	421	GAGGCCCGCGGGCCACCGGGCAGGCGTCTGCTGCTGCGGGGAGGCGCTGCTGAGGAG	480
Db	421	GAGGCCCGCGGGCCACCGGGCAGGCGTCTGCTGCTGCGGGGAGGCGCTGCTGAGGAG	480
Qy	481	GAGGCCCGGAGCTGCCCGCACGCGCTTCTGCTGCTGCTGTCATCGAGAACAGCGTCATGACC	540
Db	481	GAGGCCCGGAGCTGCCCGCACGCGCTTCTGCTGCTGCTGTCATCGAGAACAGCGTCATGACC	540
Qy	541	TCCTTCTCCCAAG	552
Db	541	TCCTTCTCCCAAG	552

```

RESULT 3
US-10-042-347-4
; Sequence 4, Application US/10042347
; Publication No. US20030114370A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Juddah
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Endostatin Protein and Peptide Fra
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 05213-0880 (43170-249874)
; CURRENT APPLICATION NUMBER: US/10/042,347
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: US 09/315,689
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 60/106,343
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 09/154,302
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 08/740,168
; PRIOR FILING DATE: 1996-10-22
; PRIOR APPLICATION NUMBER: US 60/005,835
; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-347-4

```

```

Query Match          75.3%; Score 418; DB 14; Length 546;
Best Local Similarity 85.3%; Pred. No. 1.2e-86;
Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0
Qy 1 CACACCCACGAGCACTTCAGCTGGTCTGCACCTCGTGTGGCCCTGAAACAGCCGAGCCG 60

```

1	1	CACAGGCACCGGACCTTCCAGCGGTGCTCCACTGGTTCGGCTCAACAGCCCCCTGTCA	60
61	61	GGCGCATGCGAGGCATCCGGGAGCGAGCTTCAGTGTCTTCAGAGCGCGCGCGCG	120
61	61	GGCGCATGCGGGGATCCGCGGGCCGACTTCAGTGTCTTCAGAGCGCGCGGGCCGTG	120
121	121	GGGCTGGCGGACCTTCCGGGCTTCTGTGTCGGGCTGAGGACCTCTACAGCATC	180
121	121	GGGCTGGCGGACCTTCCGGGCTTCTGTGTCCTCGCGCTGAGGACCTGTACAGCATC	180
181	181	GTGCGCGCGCGGACCGGACCGGGGTGCCGTGCTCAACCTCAGGGACGAGGTGCTCTTC	240
181	181	GTGCGCGCGGTCCGACCGGCGAGCGGTGCCCATGCTCAACCTCAGGACGAGTGTGCTTT	240
241	241	CCGAGCTGGAGGCTTATTCTCGGGTCTCCAGGGCCAGCTGAAGCCCGGGGCCGATC	300
241	241	CCGAGCTGGAGGCTTGTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGGACGATC	300
301	301	TTCTCTTTGACGGCAGAGATGTCTGTGAGCACCCCGCTGGCCCGGGAAGAGCGTGG	360
301	301	TTCTCTTTGACGGCAGGAGCGTCTGTGAGGCACCCACCTGGCCCGGGAAGAGCGTGG	360
361	361	CAGGGCTCGACCCAGCGGGGCGGCTGACGACGAGCTACTGCGAGAGCTGGCGGAGC	420
361	361	CATGGCTCGGACCCCAACGGCGCAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGAGC	420
421	421	GAGGCCCGGGGCCACCGGCGAGCGTCTGCTGCTGGCGGCGAGGCTGCTGGAGCAG	480
421	421	GAGGCTTCCCTCGGGCACCGGGCGAGGCTTCTGCTGCTGGGGGCGAGGCTCTGGGGCAG	480
481	481	GAGGCCGCGAGCTGCGGCGACGCTTCTGTTGCTCTGTCATCGAGAACGCGTCAATGACC	540
481	481	AGTGCGCGAGCTGCCATCAGCGCTACATCGTCTTGCATTGAGAACAGCTTCATGACT	540
541	541	TCCTTC	546
541	541	GCCTC	546

```

RESULT 4
US-10-292-418-3
; Sequence 3, Application US/10292418
; Publication No. US20030139365A1
; GENERAL INFORMATION:
; APPLICANT: Lo, Kin-Ming
; APPLICANT: Li, Yue
; APPLICANT: Gallies, Stephen D
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
; TITLE OF INVENTION: Immunofusins
; FILE REFERENCE: LEX-006C1
; CURRENT APPLICATION NUMBER: US/10/292,418
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/383,315
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: US 60/097,883
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(549)
; OTHER INFORMATION: endostatin
US-10-292-418-3

```

[illegible]

RESULT 5
 US-10-131-241-53
 ; Sequence 53, Application US/10131241
 ; Publication No. US20030012792A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Holaday, John W.
 ; APPLICANT: Fortier, Anne H.
 ; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
 ; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
 ; FILE REFERENCE: 05213-0344 43170-271565
 ; CURRENT APPLICATION NUMBER: US/10/131,241
 ; CURRENT FILING DATE: 2002-07-22
 ; PRIOR APPLICATION NUMBER: US 09/413,049
 ; PRIOR FILING DATE: 1999-10-06
 ; PRIOR APPLICATION NUMBER: US 09/316,802
 ; PRIOR FILING DATE: 1999-05-21
 ; PRIOR APPLICATION NUMBER: US 60/086,586
 ; PRIOR FILING DATE: 1998-05-22
 ; NUMBER OF SEQ ID NOS: 65
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 53
 ; LENGTH: 549
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-131-241-53

OS-10-292-418-3
 OTHER INFORMATION: ENDOSTATIN
 Query Match 75.3%; Score 418; DB 12; Length 549;
 Best Local Similarity 85.3%; Pred. No. 1.2e-86;
 Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
 QY 1 CACACCCACGAGACTTCCAGCTGGTGTGCACCTGGTGGCCCTGGAACAGCCGACCG 60

QY 61 GCGGCGATGCGAGGCAATCCGGGAGCGGACTTCAGTGTCTTCAGCAGGCGCGCGCGG 120
DB 61 GCGGCGATGCGGCGGCAATCCGGGAGCGGACTTCAGTGTCTTCAGCAGGCGCGCGCGG 120
QY 121 GGGTGGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTCTACAGCATC 180
DB 121 GGGTGGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTGTACAGCATC 180
QY 181 GTGGCGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTGTACAGCATC 240
DB 181 GTGGCGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTGTACAGCATC 240
QY 241 GTGGCGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTGTACAGCATC 300
DB 241 GTGGCGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTGTACAGCATC 300
QY 301 GTGGCGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTGTACAGCATC 360
DB 301 GTGGCGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTGTACAGCATC 360
QY 361 GTGGCGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTGTACAGCATC 420
DB 361 GTGGCGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTGTACAGCATC 420
QY 421 GTGGCGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTGTACAGCATC 480
DB 421 GTGGCGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTGTACAGCATC 480
QY 481 GTGGCGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTGTACAGCATC 540
DB 481 GTGGCGCGGAGCTGCCATCAGCGCTACATGCTGTGCTTGGACAGCTTCATGACT 540
QY 541 TCCTTC 546
DB 541 GCCTCC 546

RESULT 6

US-09-873-676-30
; Sequence 30, Application US/09873676
; Patent No. US2002007289A1
; GENERAL INFORMATION:
; APPLICANT: Macdonald, Nicholas J.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use
; FILE REFERENCE: 05213-0378 (43170-259333)
; CURRENT APPLICATION NUMBER: US/09/873,676
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-676-30

Query Match 75.3%; Score 418; DB 9; Length 552;
Best Local Similarity 85.3%; Pred. No. 1.2e-86;
Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 1 CACACCCACGAGGACTTCAGCTGGTGTGCTGACCTGTGGCCCTGAACAGCCCGCAGCG 60
DB 1 CACAGCCACGCGACTTCAGCCGGTGTCTCCACCTGTGTGGCTCACAGCCCTGTCA 60
QY 61 GCGGCGATGCGGCGGCACTTCGCGGCGCTTCAGTGTCTTCAGCAGGCGCGCGCG 120
DB 61 GCGGCGATGCGGCGGCACTTCGCGGCGCTTCAGTGTCTTCAGCAGGCGCGCGCG 120

QY 121 GGGTGGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTCTACAGCATC 180
DB 121 GGGTGGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTGTACAGCATC 180
QY 181 GTGGCGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTGTACAGCATC 240
DB 181 GTGGCGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCTGAGGACCTGTACAGCATC 240
QY 241 CCAGCTGGGAGGCTTCCTGTGCTGCGGCTGAGGAGCGAGTCAAGCCCGGCGCGCATC 300
DB 241 CCAGCTGGGAGGCTTCCTGTGCTGCGGCTGAGGAGCGAGTCAAGCCCGGCGCGCATC 300
QY 301 TTCTCTTTTCAGCGGAGAGATGCTCTGCGGCGGCTTCCTGTGCTGCGGCGGAGCGGTGG 360
DB 301 TTCTCTTTTCAGCGGAGAGATGCTCTGCGGCGGCTTCCTGTGCTGCGGCGGAGCGGTGG 360
QY 361 CAGGCTCCGAGCCCGGCGGCGGCTTCCTGTGCTGCGGCGGAGCGGTTCGCGGCGGAGCG 420
DB 361 CAGGCTCCGAGCCCGGCGGCGGCTTCCTGTGCTGCGGCGGAGCGGTTCGCGGCGGAGCG 420
QY 421 GAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
DB 421 GAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
QY 481 GAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
DB 481 AGTGGCGGAGCTGCCATCAGCGCTACATGCTGTGCTTGGACAGCTTCATGACT 540
QY 541 TCCTTC 546
DB 541 GCCTCC 546

RESULT 7

US-09-880-107-2178
; Sequence 2178, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2178
; LENGTH: 3394
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L22548
US-09-880-107-2178

Query Match 75.3%; Score 418; DB 10; Length 3394;
Best Local Similarity 85.3%; Pred. No. 9.2e-87;
Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 1 CACACCCACGAGGACTTCAGCTGGTGTGCTGACCTGTGGCCCTGAACAGCCCGCAGCG 60
DB 1504 CACAGCCACGCGACTTCAGCCGGTGTCTCCACCTGTGTGGCTCAACAGCCCTGTCA 1563
QY 61 GCGGCGATGCGGCGGCACTTCGCGGCGGCGGAGCTTCAGTGTCTTCAGCAGGCGCGCGCG 120
DB 1564 GCGGCGATGCGGCGGCACTTCGCGGCGGCGGAGCTTCAGTGTCTTCAGCAGGCGCGCGCG 1623
QY 121 GGGTGGCGCGGCACTTCGCGGCGCTTCCTGTGCTGCGGCGCTTCAGGACCTCTACAGCATC 180

QY 302 TCTCTTTTTCAGCGCAGAGATGTCTTCAGCAGCACCCTCGCTGGCCCGGAGAGAGCGTGTGGC 361
Db |||||
QY 301 TCTCTTTTTCAGCGCAGAGAGTCTTCAGCAGCACCCTGGCCCGGAGAGAGCGTGTGGC 360
Db |||||
QY 362 AGGCTTCGACCCCGAGCGGCGCGCTTACCGAGAGTCTTCAGCAGAGCGTGTGGCAGCG 421
Db |||||
QY 361 ATGGCTCGGACCCCGAGCGCGCAGGCTGACCGAGAGCTACTGTGAGAGCGTGTGGCAGCG 420
QY 422 AGGCTTCGAGCGCGCAGCGCGCGCGCTTCGCTGCTGGCGGCGAGGCTGTGGAGCAGG 481
Db |||||
QY 421 AGGCTTCGAGCGCGCAGCGCGCGCGCTTCGCTGCTGGCGGCGAGGCTGTGGAGCAGG 480
QY 482 AGGCTTCGAGCGCGCAGCGCGCGCGCTTCGCTGCTGGCGGCGAGGCTGTGGAGCAGG 541
Db |||||
QY 481 GTGGCGGAGCTGCCATCAGCGCTACATGCTGCTGCTGATTGAGAACAGCTTCATGACTG 540
QY 542 CCTTC 546
Db |||||
QY 541 CCTCC 545
Db |||||

RESULT 10
US-10-131-241-51
; Sequence 51, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
; LENGTH: 632
; TYPE: DNA
; ORGANISM: Murinae sp.
US-10-131-241-51

Query Match 74.9%; Score 415.6; DB 14; Length 632;
Best Local Similarity 82.8%; Pred. No. 4.1e-86;
Matches 523; Conservative 0; Mismatches 29; Indels 80; Gaps 1;
QY 1 CACACCCAGCAGACTTCAGCTGGTGTGCTGCACCTGGTGGCCCTGAACAGCCCGCAGCGG 60
Db |||||
QY 1 CATACTCATCAGGACTTTCAGCCAGTGTCTCACCTGGTGGCACTGAACACCCCTGTCT 60
Db |||||
QY 61 GCGGCGATCGAGGACTTCGCGGAGCGGACTTCAGTGTCTTCAGCAGGCGCGCGCGG 120
Db |||||
QY 61 GGAGGATCGCTGTGTATTCGTGGAGAGATTCAGTGTCTTCAGCAGGCGCGCGCGG 120
QY 121 GGCTGGCGCGGCTTCCTTCGCGGCGCTTCCTGTGCTGCGCGG----- 160
Db |||||
QY 121 GGCTGTGCGGCGACCTTCGCGGCGCTTCCTGTGCTGCTGAGGCTGCGAGTCTCTATGATC 180
QY 161 ----- 160
Db |||||
QY 181 GTGCGCGCTGTGCTACCGCGGGTCTGTGCCCATCGTCAACCTGAAGCAGAGTGTATCT 240
QY 161 TGAGGACCTCTACAGCATCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 220
Db |||||
QY 241 TGAGGACCTCTACAGCATCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 221 TCAGGAGCAGAGTGTCTTCCTCCAGCTGGAGGCGCTTATCTTCGCGGCTTCGAGGCGCAG 280
Db |||||

Db 301 TCAGGAGCAGAGTGTCTTTCAGCTGGAGGCGCTTATTTCTCGGGCTCCGAGGCGCAGC 360
QY 281 TGAAGCCCGGGCGCGCATCTTCTTTTCAGCGGAGAGATGTCTTCAGCAGCACCCTCGCT 340
Db |||||
QY 361 TGAAGCCCGGGCGCGCATCTTCTTTTCAGCGGAGAGATGTCTTCAGCAGCACCCTCGCT 420
QY 341 GSCCCCGGAAGAGCGTGTGGCAGCGCTCCGACCGCCCGGCGCGCGCTTCGACCGAGCT 400
Db |||||
QY 421 GSCCCCGGAAGAGCGTGTGGCAGCGCTCCGACCGCCCGGCGCGCGCTTCGACCGAGCT 480
QY 401 ACTGCGAGACGTGGCGGAGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTGCTG 460
Db |||||
QY 481 ACTGCGAGACGTGGCGGAGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTGCTG 540
QY 461 CGGGCAGAGCTGTGAGCAGGAGGCGCGGAGCTGCGCGCGCGCGCGCGCGCGCGCTGTG 520
Db |||||
QY 541 CGGGCAGAGCTGTGAGCAGGAGGCGCGGAGCTGCGCGCGCGCGCGCGCGCGCTGTGCTG 600
QY 521 TCGAGAACAGCGTTCATGACCTCTTCTTCCAG 552
Db |||||
QY 601 TCGAGAACAGCGTTCATGACCTCTTCTTCCAG 632
Db |||||

RESULT 11
US-10-042-347-6
; Sequence 6, Application US/10042347
; Publication No. US20030114370A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Endostatin Protein and Peptide Fra
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 05213-0880 (43170-249874)
; CURRENT APPLICATION NUMBER: US/10/042,347
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: US 09/315,689
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 60/106,343
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 09/154,302
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 08/740,168
; PRIOR FILING DATE: 1996-10-22
; PRIOR APPLICATION NUMBER: US 60/005,835
; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-347-6

Query Match 74.0%; Score 410.8; DB 14; Length 534;
Best Local Similarity 85.6%; Pred. No. 5.2e-85;
Matches 457; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 13 GACTTCCAGCTGTGCTGCACCTGTGGCCCTGAACAGCGCGCGCGCGCGCGCGCGCG 72
Db 1 GACTTCCAGCGCGTGTCTTCCACTGTGGCTTCAACAGCGCGCGCGCGCGCGCGCG 60
QY 73 GGCATCCGCGGAGCGGACTTCCAGTCTTCAGCAGGCGCGCGCGCGCGCGCGCGCG 132
Db 61 GGCATCCGCGGCGCGGACTTCCAGTCTTCAGCAGGCGCGCGCGCGCGCGCGCGCG 120
QY 133 ACCTTCCGCGCGCTTCTGTGCTGCGGGCTGCGAGGACCTTACAGCATCGTGGCGCG 192
Db 121 ACCTTCCGCGCGCTTCTGTCTCTCGCGCGCTGCGAGGACCTTACAGCATCGTGGCG 180

QY	421	GAGGCCCCGGCGCCACCGGGCAGGCGTCGTCCGCTGCTGCGGGCAGGCTGTGGAGCAG	480
Db	424	GAAGTACTTGGGGCTACAGGTCAAGGCTCTCCCTGCTGTCAAGCAGGCTCTTGGAAACAG	483
QY	481	GAGGCGCGAGCTGCCGCCACCGCTTCGTGGTGTCTGTCATCGAGAACAGCGTCATGACC	540
Db	484	AAAGCTGCCGAGCTGCCCAACAGCTACATCGTCTGTGCATTGAGAATAGCTTCATGACC	543
QY	541	TCCTTCTCCAAAGTAG	555
Db	544	TCCTTCTCCAAATAG	558

Search completed: August 17, 2003, 16:34:30
Job time : 218.125 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2003, 13:19:37 ; Search time 1617.68 Seconds
(without alignments)
8338.472 Million cell updates/sec

Title: US-09-938-391-3

Perfect score: 555

Sequence: 1 cacaccaccaggactcca.....tgacctctcttccaagtag 555

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:**

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est4:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pbg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	439.8	79.2	551	10	BF074459
2	418	75.3	707	10	BE908201
3	418	75.3	757	10	BE906253
4	418	75.3	881	14	CD105862

C	5	394.2	71.0	944	13	BU859398	AGENCOURT
C	6	391.8	70.6	657	14	CB444165	CB444165 695295 MA
C	7	384.4	69.3	703	13	BU615520	UI-H-FGO-
C	8	384.2	69.2	929	13	BQ672290	AGENCOURT
C	9	377	67.9	683	12	BM683067	UI-E-E01-
C	10	376.6	67.9	682	9	AW089583	xg20f03.x
C	11	375.8	67.7	843	10	BF385854	602046021
C	12	375.8	67.7	874	12	BI412588	602990468
C	13	374.8	67.5	715	14	CB596713	AGENCOURT
C	14	371.6	67.0	832	10	BG387051	602454749
C	15	370.2	66.7	979	13	BQ673186	AGENCOURT
C	16	369.6	66.6	897	12	BI080524	602877005
C	17	368.6	66.4	611	10	AW911243	ur83h10.Y
C	18	367.2	66.2	884	12	BI161007	602865213
C	19	366.6	66.1	1093	13	BQ723254	AGENCOURT
C	20	364.2	65.6	723	12	BI247582	602960041
C	21	363.8	65.5	753	12	BI904605	603168411
C	22	362.8	65.4	733	10	BQ072504	H311D11-
C	23	362.6	65.3	782	9	AI326391	mm18f09.x
C	24	357.4	64.4	618	9	AV696242	AV696242
C	25	350.6	63.2	664	13	BU632049	UI-H-FE1-
C	26	350.6	63.2	715	9	AUI25614	AUI25614
C	27	350.6	63.2	947	13	BU558872	AGENCOURT
C	28	348.6	62.8	720	12	BI147444	602914008
C	29	348.6	62.8	915	10	BF166139	601776586
C	30	344.6	62.1	652	13	BU352506	603527982
C	31	344.2	62.0	650	9	AI858615	w140f01.x
C	32	342.6	61.7	639	12	BM98137	UI-H-DT1-
C	33	341.4	61.5	614	13	BU459935	60367327
C	34	341	61.4	634	13	BQ772348	UI-H-EZ1-
C	35	340.4	61.3	587	9	AA288198	v515b10.f
C	36	340	61.3	660	9	AW192502	x145e10.x
C	37	339.2	61.1	690	12	BI219399	602936756
C	38	336.2	60.6	795	14	CB601615	AGENCOURT
C	39	335.2	60.4	634	9	AI970297	wf09c02.x
C	40	334.8	60.3	706	13	BU439577	603208252
C	41	333.8	60.1	745	9	AW243446	xm95d11.x
C	42	333.6	60.1	772	10	BF178591	601807962
C	43	331.6	59.7	616	13	BU158380	AGENCOURT
C	44	330.6	59.6	623	14	CB424313	598557 MA
C	45	328.2	59.1	592	13	BQ567187	gi80h10.Y

ALIGNMENTS

RESULT 1
BF074459 551 bp mRNA linear EST 25-APR-2001
LOCUS 221883 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BF074459
ACCESSION BF074459.1 GI:10867970
VERSION BF074459.1
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 551)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W., Rohrer,G.A., Chitko-McKown,C.G., Perteau,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keefe,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
PUBMED 11282978
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366

Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.

KEYWORDS	REFERENCE
SOURCE	AUTHOR
ORGANI	TITLE
	JOURNA
	COMMENT

EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 707)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.

FEATURES
source

Email: csapbbs-@email.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies
cDNA Library Arrayed by: the I.M.A.G.E. Consortium
DNA Sequencing by: Incyte Genomics
Clone distribution: MGC clone distribution
found through the I.M.A.G.E. Consortium
<http://image.llnl.gov>
Plate: L14M9704 row: d column: 16
High quality sequence stop: 688.

FEATURES
source

```

Location/Qualifiers
1. .707
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3902175"
/tissue_type="epitheloid c
/lab_host="DH10B (phage-res
/clone_lib="NIH_MGC_70"
/notes="Organ: pancreas; Vec
Site.2: Sali; Cloned undir
Average insert size 1.1 kb.
Technologies."

```

RESULT 2	BE908201	LOCUS	DEFINITION	BE908201	601500458F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3902175 5', mRNA sequence.	707 bp	linear	EST 20-OCT-2000
ACCESSION	BE908201	VERSION	BE908201.1	GI:10402537				

Db	276	CCACGTCGGAGGCTCTGTTCTCAGGCTCTGAGGCTCGGTGAAGCCCGGGACACGATC	339
Qy	301	TTCTCTTTTCAGCGCAGAGATGTCCTGACGACACCCCGCTTGGCCCCCGAAGAGCGTGTGG	360
Db	336	TTCTCTTTTTCAGCGCAAGACGTCCTGAGGCAACCCACCTTGGCCCCCAGAGAGCGTGTGG	395
Qy	361	CACGGCTCCGACCCAGCGGGCGCGCTTACCGACAGCTACTTCGAGAGCGTGGCGGACG	420
Db	396	CATGGCTCGGACCCCAACCGGGCGCAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGACG	455
Qy	421	GAGGCCCGCGCGCCACCGGCGCAGGCGTCTGCTGCTGCTGCGGCGCAGGCTGCTCGAGCAG	480
Db	456	GAGGCTCCCTCGGCGCAGCGGCGAGGCTCTCTGCTGCTGGGGGCGCAGGCTCCTGGGGCAG	515
Qy	481	GAGGCCCGCAGCTGCCCGCCACCGCTTCTGGTGGTCTCTGCAATCGAGAACAGCGTCAATGACC	540
Db	516	AGTGCCGCGAGCTGCCATCAGCGCTACATCGTCTGCAATGAGAACAGCTTTCATGACT	575
Qy	541	TCCTTC 546	
Db	576	GCCTCC 581	
RESULT 4			
CD105862			
LOCUS			
DEFINITION			
AGENCY 14021788 NIH MGC 179 Homo sapiens cDNA clone			
IMAGE:30365831 5', mRNA sequence.			
CD105862			
CD105862.1 GI:30759036			
EST.			
Homo sapiens (human)			
Homo sapiens			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 881)			
NIH-MGC http://mgc.nci.nih.gov/.			
National Institutes of Health, Mammalian Gene Collection (MGC)			
Unpublished			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-r@mail.nih.gov			
Tissue Procurement: Dr. Michael Brownstein			
cDNA Library Preparation: Invitrogen Corp			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Agencourt Bioscience Corporation			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: NDAM424 row: g column: 24			
High quality sequence stop: 689.			
Location/Qualifiers			
1..881			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:30365831"			
/tissue_type="Pituitary"			
/lab_host="DH10B-Ton A (T1 and T5 phage resistances) "			
/clone_lib="NIH_MGC_179"			
/notes="Organ: Brain; Vector: pCMV-SPORT6.1; Site1: EcoRV			
(destroyed); Site2: NotI; Library is oligo-dT primed and			
directionally cloned (EcoRV site is destroyed upon cloning			
). Average insert size 1.1 kb. Library was constructed by			
(Invitrogen). Note: this is a NIH_MGC Library."			
BASE COUNT 138 a 341 c 278 g 124 t			
ORIGIN			
Query Match 75.3%; Score 418; DB 14; Length 881;			
Best Local Similarity 85.3%; Pred. No. 3.9e-68;			
Matches 466; Conservative 0; Mismatches 80; Indels 0; Gaps 0;			
Qy	1	CACACCACGAGGACTTCAGCTGGTGGCTGCACTTGGTGGCCCTGAACAGCCCGCAGCCG	60

```
Db 292 CACAGCCACCGCGACTTCCAGCGGTGCTCCAGCTGTTGGCTCAACAGCCCCCTGTCA 351
Qy 61 GCGGCGATCGAGGATCCCGGAGCGGACTTCCAGTGTCTTCAGAGGCGCGCGCGG 120
Db 352 GCGGCGATCGGCGGATCCCGGCGGCGGACTTCCAGTGTCTTCAGAGGCGCGCGCGG 411
Qy 121 GGGTGGCGCGGCGACTTCCCGGCGGCTTCTCTGCTCGCGGTTCAGGACCTCTACAGCATC 180
Db 412 GGGTGGCGGCGACTTCCCGGCGGCTTCTCTGCTCGCGGTTCAGGACCTGTACAGCATC 471
Qy 181 GTGGCGCGCGCGAGCGACCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 472 GTGGCGCGCGCGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 531
Qy 241 CCAGCTGGGAGGCTTATCTCGGCTCGAGGCGGCTGAGGCGGCGGCGGCGGCGGCGG 300
Db 532 CCAGCTGGGAGGCTTATCTCGGCTCGAGGCGGCTGAGGCGGCGGCGGCGGCGGCGG 591
Qy 301 TTCTCTTTTCAGCGCGAGAGATGCTTTCAGGACCGCGGCTGCGGCGGCGGCGGCGGCGG 360
Db 592 TTCTCTTTTCAGCGCGAGAGATGCTTTCAGGACCGCGGCTGCGGCGGCGGCGGCGGCGG 651
Qy 361 CAGGCTTCGACCGCGGCGGCGGCTTTCAGGACCGCGGCTGCGGCGGCGGCGGCGGCGG 420
Db 652 CATGGCTCGGACCGCGGCGGCGGCTTTCAGGACCGCGGCTGCGGCGGCGGCGGCGGCGG 711
Qy 421 GAGGCGCGGCGGCGGCGGCGGCGGCTTTCAGGACCGCGGCTGCGGCGGCGGCGGCGGCGG 480
Db 712 GAGGCGCGGCGGCGGCGGCGGCGGCTTTCAGGACCGCGGCTGCGGCGGCGGCGGCGGCGG 771
Qy 481 GAGGCGCGGAGCTGCGGCGGCGGCGGCTTTCAGGACCGCGGCTGCGGCGGCGGCGGCGG 540
Db 772 AGTGGCGGAGCTGCGGCGGCGGCGGCTTTCAGGACCGCGGCTGCGGCGGCGGCGGCGG 831
Qy 541 TCCTTC 546
Db 832 GCCTCC 837
```

RESULT 5

```
BU859398/c
LOCUS
DEFINITION
IMAGE: 6650260 5', mRNA sequence.
BU859398
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 944)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: AFCC
CDNA Library Preparation: Rubin Laboratory
CDNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC2895 row: d column: 04
High quality sequence stop: 572.
```

```
Location/Qualifiers
1. 944
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6650260"
/tissue_type="teratocarcinoma, cell line"
```

FEATURES

source

```
CB444165
695295 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.
CB444165
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 657)
Smith, T.P.L., Roberts, A.J., Echtenkamp, S.E., Chitko-McKown, C.G.,
Wray, J.E. and Keesle, J.W.
```

```
/lab host="DH10B (phage-resistant)"
/clone lib="NIH_MGC_109"
/notes="organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
```

```
BASE COUNT 163 a 331 c 260 g 190 t
ORIGIN
```

```
Query Match 71.0%; Score 394.2; DB 13; Length 944;
```

```
Best Local Similarity 83.7%; Pred. No. le-63;
```

```
Matches 458; Conservative 0; Mismatches 88; Indels 1; Gaps 1;
```

```
Qy 1 CACACCCACCGAGACTTCCAGCTGGT-GGTGCACCTGTGGTGGCTTCAACAGAGCCCGAGCC 59
Db 712 CAGGTGCGACCGAGGCTTCCAGCGGTGGGTCCACATGGTTGCGCTCAGCAGCCCCCTGTC 653
Qy 60 GGGCGGATCGCAGGATCCCGGAGCGGACTTCCAGTGTCTTCAGAGGCGCGCGCGC 119
Db 652 AGCGGCGATGCGGGGCGATCCCGGCGGCGGACTTCCAGTGTCTTCAGAGGCGCGCGCG 593
Qy 120 GGGGCTGGCGGCGACCTTCCCGGCGCTTCTGTCGTCGGGCTGAGGACCTCTACAGCAT 179
Db 592 GGGGCTGGCGGCGACCTTCCCGGCGCTTCTGTCGTCGGGCTGAGGACCTCTACAGCAT 533
Qy 180 CGTGGCGCGCGCGAGCGGCGGCTGCGTCTGCTCAACCTCAGGAGCGAGGTGCTCTT 239
Db 532 CGTGGCGCGCGCGAGCGGCGGCGGCTGCGTCTGCTCAACCTCAGGAGCGAGGTGCTCTT 473
Qy 240 CCCAGCTGGGAGGCTTATCTCGGCTTCGAGGCGGCGGCTGAGGCGGCGGCGGCGG 299
Db 472 TCCAGCTGGGAGGCTTATCTCGGCTTCGAGGCTTCGAGGCGGCGGCGGCGGCGG 413
Qy 300 CTCTCTTTTCAGCGCGAGAGATGCTTCCAGCAGCCCCCGCTGCGGCGGCGGAGCGGTG 359
Db 412 CTCTCTTTTCAGCGCGAGAGATGCTTCCAGCAGCCCCCGCTGCGGCGGCGGAGCGGTG 353
Qy 360 GCAGCGCTCCGACCGCGGCGGCGGCTGCGTCAACAGAGCTACTCGAGAGCTGCGGAG 419
Db 352 GCATGGCTCGGACCGCGGCGGCGGCTGCGTCAACAGAGCTACTGTGAGAGCTGCGGAG 293
Qy 420 GGAGGCGCGGCGGCGGCGGCGGCTGCGTCTGCTGCGGCGGCGGCGGCGGCGGCGG 479
Db 292 GGAGGCTCCCTCGGCGGCGGCGGCGGCTTCTGCTGCGGCGGCGGCGGCGGCGG 233
Qy 480 GGAGGCGCGGAGCTCGCGCGGCGGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
Db 232 GAGTCCCGGAGCTCCGCTACGCGCTACATGCTGCTCTGCTGCTGCTGCTGCTGCTG 173
Qy 540 CTCCTTC 546
Db 172 TGCTTC 166
```

RESULT 6

CB444165/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 657)

REFERENCE

AUTHORS

Seq primer: M13 Forward
POLYA=Yes.

```

FEATURES
source
Location/Qualifiers
1. .683
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-E01-ajd-g-02-01"
/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-E01"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-E01 is a normalized cDNA library containing the
following tissue(s): fetal eye. The library was
constructed according to Bonaïdo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
CGCGTATACC. This library was created for the program, Gene
Discovery in the Visual System, supported by National Eye
Institute (NEI).
TAG LIB=UI-E-E01
TAG_TISSUE=human fetal eye
TAG_SEQ=CGCGTATACC"

```

```

442 CTGAAGCCGGGACGCTCTTCTCTTTGACGCAAGGACGCTCTGAGGACACCCACC 383
340 TGGCCCGGAGAGCGTGTGGACGCTCGACCCAGCGGGCGCGCTGACCGACAGC 399
382 TGGCCCGGAGAGCGTGTGGACGCTCGACCCAGCGGGCGCGCTGACCGACAGC 323
400 TACTGCGAGACGTTGGCGGACGAGGCGCCCGGGCGGACCGGGCAGGCGTCTGCTG 459
322 TACTGTGAGACGTTGGCGGACGAGGCTCCTCGGCCACGGCCAGGCTCTGCTGCTG 263
460 GCGGGCAGGCTGTGAGCAGGAGGCGCGAGTGTGCGGCACCGCTTGTGTGCTCTGC 519
262 GCGGGCAGGCTCCTGCGGACGAGTGTGCGGAGTGTGCGGACGCTACATCGTGTCTGC 203
520 ATCGAGACAGCGTCTATGACCTCTTC 546
202 ATTGAGAACAGCTTTCATGCTGCTCC 176

RESULT 11
BF385854
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BF385854 843 bp mRNA linear EST 27-NOV-2000
602046021F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:4195660 5',
mRNA sequence.
BF385854
BF385854.1 GI:11367159
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 843)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9531 row: i column: 05
High quality sequence stop: 761.
Location/Qualifiers
1..843
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4195660"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Li9"
/note="Organ: liver; Vector: pCMV-SPORTs; Site 1: Not I;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 167 a 278 c 228 g 170 t
ORIGIN

Query Match 67.7%; Score 375.8; DB 10; Length 843;
Best Local Similarity 79.8%; Pred. No. 2.6e-60;
Matches 443; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 1 CACACCCACGAGGACTTCCAGCTGGTGTGTCACCTGGTGGCCCTGAACAGCCGCGAGCG 60
Db 171 CATACTCATCAGGACTTTGACGAGTGTCTCCACTGTGTGGCATGACACCCCTGTCT 230
QY 61 GCGCGCATGCGAGGCAATCCGGGAGCGGAGCTTCAGTGTCTTCAGAGCGCGCGCG 120

```

```

231 GGAGGACATCGCTGGTATCCGTGGAGCAGATTTCAGTGCTTCCAGAACCCCGACCGTG 290
121 GGGCTGGCGGCGCACCTTCCGGGCTTCTGTCGTGCGGCTGCGAGACTCTTACAGCATC 180
291 GGGCTGTGCGGACCTTCCGGGCTTCTGTCCTTAGCTGCGAGATCTCTATAGCATC 350
181 GTGCGCCCGCGCGACCCGACCCGGGTGCGCTGCTCAACCTCAGGACGAGGTCTCTTC 240
351 GTGCGCCGTGCTGACCCGGGGTCTGTGCCATCGTCAACCTGAAGGACGAGGTCTATCT 410
241 CCCAGCTGGAGGCTTATTCTGGGCTCCGAGGCCAGCTGAAGCCCGGGGCCCGCATC 300
411 CCCAGCTGGGACTCCCTGTCTTCTGCTCCAGGGTCAACTGCAACCCGGGGGCCCGCATC 470
301 TTCTCTTCGACGCGCAGAGATGCTCTGCGACACCCCGCTGCGCCCGGAGAGCGTGG 360
471 TTTCTTTTGACGCGCAGAGATGCTCTGAGACACCCAGCTGSCCGCAGAGCGTATGG 530
361 CACGCTCCGACCCCGCGCGCTGACCCGACAGTACTGCGAGACGTGGCGGAGC 420
531 CACGCTCCGACCCCGCGCGCTGATGGAGAGTTACTGTGAGACATGGCGAACT 590
421 GAGGCCCGCGGCGCACCCGGGCGAGCGTCTGCTGCTGGCGGCGAGCTGCTGAGCAG 480
591 GAAACTACTGCGGCTTACAGGTACAGCTCCTCCCTGCTGTACAGGAGCTCCTGGAACAG 650
481 GAGGCCGCGAGCTGCGCGCACCCCTTCTGTTGGTGTCTGTCATCGAGAACAGCTCATGACC 540
651 AAAGCTGCGAGCTGCCAACAGCTACATGCTCTGTCATTTGAGATAGTTCATGACC 710
541 TCCTTCTCCAAAGTAG 555
711 TCTTCTCCAAATAG 725

RESULT 12
BI412588
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BI412588 874 bp mRNA linear EST 14-AUG-2001
602990468F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5146409 5',
mRNA sequence.
BI412588
BI412588.1 GI:15173511
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 874)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11361 row: g column: 18
High quality sequence start: 23
High quality sequence stop: 808.
Location/Qualifiers
1..874
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:5146409"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"

```



```

QY 542 CCTTCTCCAAGTAG 555
Db 543 CTTTCTCCAATAG 556

RESULT 14
BG387051
LOCUS 602454749F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4582933 5',
DEFINITION mRNA sequence.
ACCESSION BG387051
VERSION BG387051.1 GI:13280497
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1306 row: a column: 14
High quality sequence stop: 679.
FEATURES
source
Location/Qualifiers
1..832
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4582933"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 15"
/notes="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT 129 a 320 c 265 g 118 t
ORIGIN
Query Match 67.0%; Score 371.6; DB 10; Length 832;
Best Local Similarity 83.6%; Pred. No. 1.6e-59;
Matches 445; Conservative 0; Mismatches 84; Indels 3; Gaps 2;

QY 1 CACACCCACGAGACTTCACGTGGTGTGTCACCTGAGTGGCCCTGAACAGCCGCGCG 60
Db 214 CACAGCCACCGGACATCCCGGGCGGCTCCACCTGGTGGCTCAACAGCCCTGTCA 273
QY 61 GCGCGCATCGAGGCATCCGGGAGCGGACTTCAGTGTCTTCAGCAGCGCGCGCGCG 120
Db 274 GCGCGCATCGGGGACATCCCGGGCGGCTCCAGTGTCTTCAGCAGCGCGCGCG 333
QY 121 GGGCTGGCGGACACTTCGGGGCTTCCTGTCGTGCGGGCTGAGACCTCTACAGCATC 180
Db 334 GGGCTGGCGGACACTTCGGGGCTTCCTGTCGTGCGGGCTGAGACCTGTACAGCATC 393
QY 181 GTGCGCGGCGGACCGACCGGCGGTCGTGTCGTCGTCGTCGTCGTCGTCGTCGTC 240
Db 394 GTGCGCGGCGGACCGGCGGCGGTCGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 453

```

```

QY 241 CCAGCTCGGAGGCTTATTCTCGGGCTCCGAGGGCCAGCTGAAGCCCGGGGCCCGCATC 300
Db 454 CCCAGCTGGAGGCTCTCTTCTCAGGCTCTGAGGTCCTGAGCCCGGGGCAAGCATC 513
QY 301 TTCTCTTTGAGCGGAGAGATGCTCTGAGCAGCACCCGCTGCGCCCGGAGAGCGTGTGG 360
Db 514 TTCTCTTTGAGCGGAGAGATGCTCTGAGCAGCACCCGCTGCGCCCGGAGAGCGTGTGG 573
QY 361 CACGGCTCCGACCCCGAGCGGGCGCC -GCCTGACCGACAGCTACTGCGAGAGCTGGCGGA 418
Db 574 CATGGCTCGACCCCAACGGGCGGAGGCTGACCGAGAGCTACTGTGAGAGCTGGCGGA 633
QY 419 CGAGGCGCCCGCGGCCACCGGGAGGCTGCTGCTGCTGCGGGGAGGCTGCTGGAGC 478
Db 634 CGAGGCTCCCTCGGACACGGCCAGGCTC -TCGCTGCTGGGGGCGAGGCTCCTGGGCG 692
QY 479 AGGAGCGCGAGCTGCGGCCACCGCTTCGTGGTCTCTGCAATCGAGACAG 530
Db 693 AGAGTGCCGAGAGCTTGCTTCACGCCCTCATCGTCTGTCATTGAGAACAG 744

RESULT 15
BG673186
LOCUS AGENCOURT_8414390 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6272287
DEFINITION 5', mRNA sequence.
ACCESSION BG673186
VERSION BG673186.1 GI:21784020
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2450 row: g column: 08
High quality sequence stop: 638.
FEATURES
source
Location/Qualifiers
1..979
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6272287"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 102"
/notes="Organ: salivary gland; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 171 a 363 c 301 g 144 t
ORIGIN
Query Match 66.7%; Score 370.2; DB 13; Length 979;
Best Local Similarity 83.1%; Pred. No. 2.9e-59;
Matches 458; Conservative 0; Mismatches 88; Indels 5; Gaps 3;

QY 1 CACACCCACGAGACTTCACGTGGTGTGTCACCTGAGTGGCCCTGAACAGCCGCGCG 60

```


Db 227 CACAGCCACCGGACTTCACGCGGTCTCCACCTGGTTGCGCTCAACAGCCCCCTGTCA 286
Qy 61 GCGCGCATCGGAGGATCGGGAGCGGACTTCCAGTGTCTCCAGCAGCGCGCGCGG 120
Db 287 GCGCGCATCGGGGATCCCGGGGCGGACTTCCAGTGTCTCCAGCAGCGCGGGCGG 346
Qy 121 GGGCTGGCGGCGACCTTCGCGGCTTCTGTGTCGCGGCTGCGGACCTCTACAGCATC 180
Db 347 GGGCTGGCGGCGACCTTCGCGGCTTCTGTGTCGCGGCTGCGGACCTGTACAGCATC 406
Qy 181 GTGGCGGCGGCGACCGGACCGGGGTGCGGCTGCTCAACCTCAGGACGAGGTGCTTTC 240
Db 407 GTGCGCGGTGCGGACCGGCGGCGGCTGCGGCTGCTCAACCTCAGGACGAGGTGCTTTC 466
Qy 241 CCCAGCTGGAGGCTTATTTCTCGGCTCCGAGGCGGCTGAGCGGCGGCGGCGGCGCATC 300
Db 467 CCCAGCTGGAGGCTGTTCTCAGGCTCTGAGGCTCGGCTGAGGCGGCGGCGGCGCATC 526
Qy 301 TTCTCTTTTCGACGCGAGATGTCTGACGACCGGCTGCGGCGGCGGCGGCGGCGGCGG 360
Db 527 TTCTCTTTTAAAGGCAAGGAGCTCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 586
Qy 361 CAGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
Db 587 CATGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 646
Qy 421 GAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 479
Db 647 GAGGCTCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 706
Qy 480 GGA--GGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 535
Db 707 ACAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 766
Qy 536 TGACCTCTTTC 546
Db 767 TGAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 777

Search completed: August 17, 2003, 16:22:58
Job time : 1619.68 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 17, 2003, 18:39:24 ; Search time 2928.89 Seconds
(without alignments)
2570.043 Million cell updates/sec

Title: US-09-938-391-4
Perfect score: 966
Sequence: 1 HTHQDFQLVHLVALNSQP.....CRHAFVVLCIENSVMTSFSK 184

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09938391/runat_04082003.130656.27089/app_query.fasta.1.718
-DB=GenEmbl -QFMT=fastap -SURFIX=p2n.rge -MINWATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938391@cgn 1 1.5066 @runat_04082003.130656.27089 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	966	100.0	555	6	AX399631	Sequence
2	966	100.0	829	6	AX399629	Sequence
3	835	86.4	546	6	AR193165	Sequence
4	835	86.4	549	6	AX100086	Sequence
5	835	86.4	552	6	AX395662	Sequence
6	835	86.4	552	6	AX473835	Sequence
7	835	86.4	552	6	BD081407	Fused pro
8	835	86.4	564	9	AF416592	Homo sapi
9	835	86.4	786	9	AF282883	Homo sapi
10	835	86.4	900	6	AX370851	Sequence
11	835	86.4	2364	9	AK098216	Homo sapi
12	835	86.4	2637	9	BC033715	Homo sapi
13	835	86.4	3394	6	AX409531	Sequence
14	835	86.4	3394	6	IS1045	Sequence
15	835	86.4	3394	9	HUMCOL18AX	Human colla
16	835	86.4	5408	9	AF018082	Homo sapi
17	835	86.4	5929	9	AF018081	Homo sapi
18	831	86.0	552	10	AF257775	Mus muscu
19	831	86.0	555	6	BD102800	Process f
20	831	86.0	555	6	BD128313	Endostati
21	831	86.0	558	6	AR268609	Sequence
22	831	86.0	565	6	AR100642	Sequence
23	831	86.0	565	6	E34073	Carrier/DNA
24	831	86.0	573	6	AR183471	Sequence
25	831	86.0	573	6	AR183520	Sequence
26	831	86.0	573	6	AR217279	Sequence
27	831	86.0	573	6	AR282751	Sequence
28	831	86.0	573	6	AX042272	Sequence
29	831	86.0	573	6	AX490666	Sequence
30	831	86.0	624	6	AX084406	Sequence
31	831	86.0	1615	10	MMU03714	Mus musculu
32	831	86.0	4437	10	MUSCOLLAG	Mus musculu
33	831	86.0	4802	10	BC043697	Sequence
34	828	85.7	624	6	AX128409	Sequence
35	828	85.7	4031	6	IS1044	Sequence
36	828	85.7	4031	10	MUSCOLPRO	Mouse alpha
37	822	85.1	555	9	AF184060	Homo sapi
38	817	84.6	534	6	AR193166	Sequence
39	817	84.6	537	6	AX100092	Sequence
40	813	84.2	851	10	AF189709	Rattus no
41	799	82.7	650	6	AX370853	Sequence
42	754	78.1	515	10	RNO236873	Rattus no
43	750	77.6	5279	5	AF083440	Gallus ga
44	694	71.8	5598	5	AB047066	Xenopus 1
45	691	71.5	5358	5	AY052763	Xenopus 1

ALIGNMENTS

AX399631
LOCUS AX399631 555 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 3 from Patent EP1191036.
ACCESSION AX399631
VERSION AX399631.1 GI:21335410
KEYWORDS
SOURCE
ORGANISM
Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1
REFERENCE
AUTHORS Sheppard, M.G. and Tong, X.
TITLE Methods and compositions for diagnosing and treating disorders involving angiogenesis
JOURNAL Patent: EP 1191036-A 3 27-MAR-2002;
Pfizer Products Inc. (US)
FEATURES
source
Location/Qualifiers
1..555
/organism="Canis familiaris"
/mol_type="genomic DNA"
/db_xref="taxon:9615" 83 t
BASE COUNT 78 a 203 c 191 g 83 t
ORIGIN
Alignment Scores:
Pred. No.: 1,096-69 Length: 555
Score: 966.00 Matches: 184
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-938-391-4 (1-184) x AX399631 (1-555)
QY 1 HistHrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
DB 1 CACACCCACAGGACTTCCAGCTGGTGTGACCTGGTGGCCCTGAACAGCGCGAGCGG 60
QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
DB 61 GCGGGATCGAGGATCCGGGAGGAGGACTTCCAGTGTTCAGCAGGCGCGCGCGG 120
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrsIle 60
DB 121 GGGCTGGCGGACCTTCCGGGCTTCTGCTGCTGGCTGAGACCTCTACAGCATC 180
QY 61 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 181 GTGCGCGCGCGACCGGCGTGGCGTGGCTGCTCAACCTCAGGAGCAGGTGCTTTC 240
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 100
DB 241 CCAGCTGGAGGCTTATTTCTGGGCTCCGAGGCGCAGTGAAGCCCGGGCGCGCATC 300
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
DB 301 TTCTCTTCGAGCGCAGAGATGCTTCGAGCACCCTGGCCCTGGCCCGGAGAGCGTGTG 360
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrcysGluThrTrpArgThr 140
DB 361 CACGGCTCCGACCCAGCGGCGCGCTGACCGACAGCTACTGCGAGACGTGGCGGACG 420
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuAlaGlyArgLeuGluGln 160
DB 421 GAGGCCCGCGCGCCACCGGCGAGGCGTCTGCTGCTGGCGGCGAGGCTGTGGAGCAG 480
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
DB 481 GAGGCCCGGAGTCCGCCACCGCTCTGCTGGTGTCTCTGATCGAGACAGCGTCATGACC 540
QY 181 SerPheSerLys 184
DB 541 TCCTTCTCCAAG 552

RESULT 2
LOCUS AX399629 829 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 1 from Patent EP1191036.
ACCESSION AX399629
VERSION AX399629.1 GI:21335409
KEYWORDS
SOURCE
ORGANISM
Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1
REFERENCE
AUTHORS Sheppard, M.G. and Tong, X.
TITLE Methods and compositions for diagnosing and treating disorders involving angiogenesis
JOURNAL Patent: EP 1191036-A 1 27-MAR-2002;
Pfizer Products Inc. (US)
FEATURES
source
Location/Qualifiers
1..829
/organism="Canis familiaris"
/mol_type="genomic DNA"
/db_xref="taxon:9615" 113 t
BASE COUNT 124 a 314 c 278 g 113 t
ORIGIN
Alignment Scores:
Pred. No.: 1,696-69 Length: 829
Score: 966.00 Matches: 184
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-938-391-4 (1-184) x AX399629 (1-829)
QY 1 HistHrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
DB 139 CACACCCACAGGACTTCCAGCTGGTGTGACCTGGTGGCCCTGAACAGCGCGAGCGG 198
QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
DB 199 GCGCGCATCGAGGATCCGGGAGGAGGACTTCCAGTGTTCAGCAGGCGCGCGCGG 258
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrsIle 60
DB 259 GGGCTGGCGGACCTTCCGGGCTTCTGCTGCTGGCTGAGACCTCTACAGCATC 318
QY 61 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 319 GTGCGCGCGCGACCGGCGTGGCGTGGCTGCTCAACCTCAGGAGCAGGTGCTTTC 378
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 100
DB 379 CCAGCTGGAGGCTTATTTCTGGGCTCCGAGGCGCAGTGAAGCCCGGGCGCGCATC 438
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
DB 439 TTCTCTTCGAGCGCAGAGATGCTTCGAGCACCCTGGCCCTGGCCCGGAGAGCGTGTG 498
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrcysGluThrTrpArgThr 140
DB 499 CACGGCTCCGACCCAGCGGCGCGCTGACCGACAGCTACTGCGAGACGTGGCGGAGG 558
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuAlaGlyArgLeuGluGln 160
DB 559 GAGGCCCGCGGCGCACCGGCGAGGCGTCTGCTGCTGGCGGCGAGGCTGTGGAGCAG 618
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
DB 619 GAGGCCCGGAGTCCGCCACCGCTCTGCTGGTGTCTCTGATCGAGACAGCGTCATGACC 678
QY 181 SerPheSerLys 184


```
Db 541 GCC 543
RESULT 5
LOCUS AX395662 552 bp DNA linear PAT 18-MAY-2002
DEFINITION Sequence 30 from Patent WO0193897.
ACCESSION AX395662
VERSION AX395662.1 GI:21066471
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Sim, K.L. and Macdonald, N.J.
AUTHORS Angiostatin and endostatin binding proteins and methods of use
TITLE Patent: WO 0193897-A 30 13-DEC-2001;
JOURNAL Entremed, Inc. (US)
FEATURES
Location/Qualifiers
1..552
Source /organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 83 a 196 c 179 g 94 t
ORIGIN
Alignment Scores:
Pred. No.: 4.3e-59 Length: 552
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: 6 Gaps: 0
US-09-938-391-4 (1-184) x AX395662 (1-552)
QY 1 HisThrHisGlnAspPheGlnLeuValHisLeuValAlaLeuAsnSerProGlnPro 20
DB 1 CACAGCCACGGGACTTCCAGCGGTGCTCCACCTGGTGGCTCAACAGCCCCCTGTCA 60
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
DB 61 GCGCGCATGCGGGCATCCGCGGGCCGACTTCCAGTGTCTCCAGCAGCGCGGGCGGTG 120
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
DB 121 GGGCTGGCGGGCACTTCCGCGCTTCTGCTCCGCGCTGTCAGACCTGTACAGCATC 180
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 181 GTGCGCGGTGCGGACCGCGCGGTGCTCCATCGTCAACCTCAAGACGAGCTGTGTTT 240
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 100
DB 241 CCCAGCTGGAGGCTCTGTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGGACGCATC 300
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
DB 301 TTCTCTTTGACGGCAAGGACGCTCTGAGGACACCCACCTCGGCCCCAGAGAGCGGTGG 360
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
DB 361 CATGGCTCGGACCCCAACGGGCGCAGCTGACCGAGAGCTACTGTGAGACGTGGCGGAGC 420
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGluGln 160
DB 421 GAGGCTCCCTCGGCCACCGGCCAGGCTCTCTGCTGCTGGGGGGCAGGCTCTCTGGGGCAG 480
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
DB 481 AGTGCGCGGAGCTGCCATCAGCGCTCATCGTGTCTGTGATTGAGAACAGCTTCATGACT 540

QY 181 Ser 181
Db 541 GCC 543
RESULT 6
LOCUS AX473835 552 bp DNA linear PAT 09-AUG-2002
DEFINITION Sequence 2 from Patent WO0230982.
ACCESSION AX473835
VERSION AX473835.1 GI:22208005
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Sim, K.L. and Liang, H.
AUTHORS Angiogenesis-inhibiting peptides and proteins and methods of use
TITLE Patent: WO 0230982-A 2 18-APR-2002;
JOURNAL Entremed, Inc. (US)
FEATURES
Location/Qualifiers
1..552
Source /organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 83 a 196 c 179 g 94 t
ORIGIN
Alignment Scores:
Pred. No.: 4.3e-59 Length: 552
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: 6 Gaps: 0
US-09-938-391-4 (1-184) x AX473835 (1-552)
QY 1 HisThrHisGlnAspPheGlnLeuValHisLeuValAlaLeuAsnSerProGlnPro 20
DB 1 CACAGCCACGGGACTTCCAGCGGTGCTCCACCTGGTGGCTCAACAGCCCCCTGTCA 60
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
DB 61 GCGCGCATGCGGGCATCCGCGGGCCGACTTCCAGTGTCTCCAGCAGCGCGGGCGGTG 120
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
DB 121 GGGCTGGCGGGCACTTCCGCGCTTCTGCTCCGCGCTGTCAGACCTGTACAGCATC 180
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 181 GTGCGCGGTGCGGACCGCGCGGTGCTCCATCGTCAACCTCAAGACGAGCTGTGTTT 240
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 100
DB 241 CCCAGCTGGAGGCTCTGTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGGACGCATC 300
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
DB 301 TTCTCTTTGACGGCAAGGACGCTCTGAGGACACCCACCTCGGCCCCAGAGAGCGGTGG 360
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
DB 361 CATGGCTCGGACCCCAACGGGCGCAGCTGACCGAGAGCTACTGTGAGACGTGGCGGAGC 420
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGluGln 160
DB 421 GAGGCTCCCTCGGCCACCGGCCAGGCTCTCTGCTGCTGGGGGGCAGGCTCTCTGGGGCAG 480
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
DB 481 AGTGCGCGGAGCTGCCATCAGCGCTCATCGTGTCTGTGATTGAGAACAGCTTCATGACT 540
```

```

QY 181 Ser 181
Db 541 GCC 543

RESULT 7
BD081407
LOCUS
DEFINITION Fused protein containing angiotensin component and utilization
ACCESSION BD081407
VERSION JP 2001518304-A/50.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Bolanowski, M.A., Caparon, M.H., Casperson, G.F., Gregory, S.A.,
Fused protein containing angiotensin component and utilization
JOURNAL Patent: JP 2001518304-A 50 16-OCT-2001;
GD SEARLE AND CO
COMMENT OS Homo sapiens (human)
PN JP 2001518304-A/50
PD 16-OCT-2001
PF 30-SEP-1998 JP 2000513958
PI 01-OCT-1997 US 60/060609
PI MARK A BOLANOWSKI, WAIRE H CAPARON, GERALD F CASPERSON, SUSAN A
PI GREGORY,
PI BARBARA K KLEIN, JOHN P MCKEARN
PC C12N15/002, A61K38/00, A61K48/00, A61P9/10, A61P35/00, C07K14/52, PC
C07K14/56
PC C07K14/78, C12N9/68, C12N15/00, A61K37/02
CC Fused protein containing angiotensin component and utilization

CC thereof in
CC antitumor therapy
FH Key Location/Qualifiers
FT source 1. .552 /organism='Homo sapiens (human)',
FT Location/Qualifiers
source 1. .552
/mol_type='genomic DNA'
/db_xref='taxon:9606'
BASE COUNT 83 a 196 c 179 g 94 t
ORIGIN

Alignment Scores:
Pred. No.: 4 3e-59 Length: 552
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: 6 Gaps: 0

US-09-938-391-4 (1-184) x BD081407 (1-552)

QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
Db 1 CACAGCCACCGGACATCCAGCGGGTCTCCACCTGTTCCGCTCAACAGCCCTGTCA 60
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
Db 61 GCGCGGATCGGGGCGATCCCGGGCGGAGTTCAGTGTTCAGCAGCGCGGGCGGTG 120
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 121 GGGCTCGGGGACCTTCGCGGCTTCCTGCTCCGCGCTGCGAGACCTGACGATC 180
QY 61 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80

```

```

Db 181 GTGCGCCGTGCGGACCGCGGACCGCTGCCATCGTCAACCTCAAGGACGAGCTGCTTT 240
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLeuProGlyAlaArgIle 100
Db 241 CCCAGCTGGAGGCTCTGTTCTCAGGCTCTGAGGCTCCGCTGAGCCCGGGGACGATC 300
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 301 TTCTCTTTGACGGCAAGGACGCTCTGAGGCAACCCACCTGCGCCCAAGAGCGGTGG 360
QY 121 HisGlySerAspProSerGlyArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
Db 361 CATGGCTCGGACCCCAACGGCGGAGGCTGACCGAGAGCTACTGTGAGACGTTGGGAGC 420
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
Db 421 GAGGCTCTCTCGGCCACGGGCCAGGCTCTCGCTGCTGGGGGGGAGGCTCTCTGGGGCAG 480
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
Db 481 AGTGGCGGAGCTGCCATCAGCCTACATCGTGTCTGCTGATTGAGAACAGCTTCATGACT 540
QY 181 Ser 181
Db 541 GCC 543

RESULT 8
AF416592
LOCUS Homo sapiens collagen XVIII mRNA, partial cds.
DEFINITION AF416592
ACCESSION AF416592
VERSION AF416592.1 GI:17226297
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Feng, Y., Qui, L.B., Liu, C.X. and Ma, Q.J.
Inhibition effect in vitro of purified endostatin expressed in
Pichia pastoris
JOURNAL Sheng Wu Gong Cheng Xue Bao 17 (3), 278-282 (2001)
MEDLINE 21409408
PUBMED 11517600
REFERENCE 2 (bases 1 to 564)
AUTHORS Feng, Y., Wu, Y., Zhu, X., Liu, C.X. and Ma, Q.J.
TITLE Endostatin contributes to maintain cartilage homeostasis via
promotion of the anabolic program of chondrocytes
JOURNAL Unpublished
AUTHORS Feng, Y., Qui, L.B. and Ma, Q.J.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-2001) Genetic Engineering, Beijing Institute of
Biotechnology, Taiping Road, Beijing 100850, P.R. China
FEATURES
Location/Qualifiers
1. .564
/organism='Homo sapiens'
/mol_type='mRNA'
/db_xref='taxon:9606'
/chromosome='21'
<1. .564
/Note='similar to endostatin; can inhibit endothelial cell
proliferation'
/codon_start=1
/product='collagen XVIII'
/protein_id='AAL37720.1'
/db_xref='GI:17226298'
/translation='SPPAHSRDFDPVLHLVALNSPLSGMRGIRGADFQCFQARAV
GLACTPRAFLSRQLDQLSIVVRADRAAVPIVNLKDELFPSEALFSGSEGLPKGA
RIFSPDGKDLVRHPTWPKQSVWHGSDPPRRRLTESYCTWRTAFSATQASSLLGGR
LLGQSAASCHAYIVLCIENSPMTASK'
misc_feature <1. .561

```

/note="Region: NC1 domain"		86 a		203 c		180 g		95 t	
BASE COUNT		86 a		203 c		180 g		95 t	
ORIGIN		86 a		203 c		180 g		95 t	
Alignment Scores:		4.41e-59		Length:		564			
Pred. No.:		835.00		Matches:		154			
Score:		92.82%		Conservative:		14			
Percent Similarity:		92.82%		Mismatches:		13			
Best Local Similarity:		85.08%		Indels:		0			
Query Match:		86.44%		Gaps:		0			
DB:		9							
US-09-938-391-4 (1-184) x AF416592 (1-564)									
QY	1	HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro	20						
Db	13	CACAGTACCGCGGACCTCCAGCGGTGCTCCACTGGTTCACAGCCCTGCTCA	72						
QY	21	GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAla	40						
Db	73	GGCGGATCGGGGATCCCGGGGCGGACTTCCAGTGTCTCCAGCAGGGCGGCGTG	132						
QY	41	GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSer	60						
Db	133	GGGCTGGGGGACCTTCGCGGCTTCTCTCTCGCGCTGCGGCTGAGGACCTGTAC	192						
QY	61	ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeu	80						
Db	193	GTGCGCGTGCAGACCGCGGCGGCGGCTCCAGTGTCTCCAGTGTCTCCAGTGT	252						
QY	81	ProSerTrpGluAlaLeuPheSerGlySerGlyGlnGlnLeuLysProGlyAlaArg	100						
Db	253	CCAGCTGGGAGGCTGTCTTCAGGCTCTGAGGCTGAGGCTGAGGCTGAGGCTG	312						
QY	101	PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerVal	120						
Db	313	TTCTCTTTGACGGCAGGACGCTCTGAGGACACCCACCTGGGCGGCGGCGGCG	372						
QY	121	HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArg	140						
Db	373	CATGGCTCGGACCCCAACGCGGCGGCGGCTGACGAGAGCTACTGTGAGACGT	432						
QY	141	GluAlaProAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln	160						
Db	433	GAGGCTCCCTCGGCGGCGGCGGCTCTCTCTGCTGGGGGCGAGGCTCTCTGG	492						
QY	161	GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMet	180						
Db	493	AGTGGCGGAGCTGCCATCAGCCTACATCGTCTCTGCTCTGCAATTCAGAACAG	552						
QY	181	Ser 181							
Db	553	GCC 555							
RESULT 9		AF282883		786 bp		mRNA		linear	
LOCUS									
DEFINITION									
ACCESSION									
VERSION									
KEYWORDS									
SOURCE									
ORGANISM									
REFERENCE									
AUTHORS									
TITLE									
JOURNAL									
FEATURES									
:source									


```

SOURCE      synthetic construct
ORGANISM    synthetic construct
            artificial sequences.
REFERENCE   1
AUTHORS     Chapman, P.W., de Luca, G. and Falciola, L.
TITLE       Method of producing functional protein domains
JOURNAL     Patent: WO 0210372-A 12 07-FEB-2002;
            Applied Research Systems ARS Holding S.A. (AN)
            Location/Qualifiers
FEATURES   source
            1..900
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
            /note="mRNA for mGSP fused to COL18A1 exons 38-41 for
            pEnd-HR41"
BASE COUNT 146 a 336 c 274 g 144 t
ORIGIN
Alignment Scores:
Pred. No.: 7,35e-59 Length: 900
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: 6 Gaps: 0
US-09-938-391-4 (1-184) x AX370851 (1-900)
QY 1 HisThrHisGlnAspPheGlnLeuValHisLeuValAlaLeuAsnSerProGlnPro 20
DB 337 CACAGCCACCGCGACTTCACGCGGTGCTCCACCTGGTTCGCTCAACAGCCCCCTGTCA 396
QY 21 GlyGlyMetArgGlyLeuAlaAspPheGlnCysPheGlnGlnAlaArgAlaAla 40
DB 397 GCGCGCATGCGGGGATCCCGGGGGCGGACTTCAGTGTCTTCAGCAGCGCGGCCGTG 456
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
DB 457 GGGCTGGGGGACCTTCGCGCCCTTCCTGTCTCGCGCTCGCAGGACCTGTACAGCATC 516
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 517 GTGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 576
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGluGlnLeuValProGlyAlaArgIle 100
DB 577 CCAGCTGGAGGCTCTGTCTCAGGCTCTGAGGTCCTGAGGTCCTGAGGTCCTGAGGTC 636
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
DB 637 TTCTCTTTTTCGCGCAAGGACGCTCTGAGGACCCACCTGCGCCCGCCAGAGAGCGTGTG 696
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
DB 697 CATGGCTGGAGCCCAACAGCGCGCGAGCTGACCGAGAGCTACTGTGACAGCGTGGCGGAG 756
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
DB 757 GAGGCTCCCTCGGCGCACGGGCGAGGCTCTCTCGTCTGGGGGGGAGGCTCTCTGGGGGAG 816
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
DB 817 AGTGCGCGGAGTGCCATCACGCTTACATCGTCTGTGATTGAGACAGCTTCATGACT 876
QY 181 Ser 181
DB 877 GCC 879
RESULT 11
LOCUS      AK098216
DEFINITION Homo sapiens cDNA FLJ40897 fis, clone UTERU2002964, highly similar
            to Homo sapiens type XVIII collagen (COL18A1) mRNA.
ACCESSION  AK098216
VERSION     AK098216.1 GI:21758185
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K.,
Arita, M., Musashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B.,
Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and
Isogai, T.
NEDO human cDNA sequencing project
Unpublished
REFERENCE   2 (bases 1 to 2364)
AUTHORS     Isogai, T. and Yamamoto, J.
TITLE       Direct Submission
JOURNAL     Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
            (E-mail: genomics@nri.co.jp, tel: 81-438-52-3975, Fax: 81-438-52-3986)
            NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
            Location/Qualifiers
            1..2364
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="UTERU2002964"
            /tissue_type="uterus"
            /clone_lib="UTERU2"
            /note="cloning vector: pME18SFL3"
BASE COUNT 420 a 768 c 776 g 400 t
ORIGIN
Alignment Scores:
Pred. No.: 2,12e-58 Length: 2364
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: 9 Gaps: 0
US-09-938-391-4 (1-184) x AK098216 (1-2364)
QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
DB 1644 CACAGCCACCGCGACTTCACGCGGTGCTCCACCTGGTTCGCTCAACAGCCCCCTGTCA 1703
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaAla 40
DB 1704 GCGCGCATGCGGGGATCCCGGGGGCGGACTTCACAGTGTCTCCAGCAGCGCGGCCGTG 1763
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
DB 1764 GGGCTGGGGGACCTTCGCGGCTCTCTGTCTCGGCTTCAGGCTTCAGGCTTCAGGCTTC 1823
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 1824 GTGCGCGGTGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1883
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGluGlnLeuValProGlyAlaArgIle 100
DB 1884 CCCAGCTGGAGGCTCTGTCTTCTCAGGCTCTGAGGCTCCGCTGAAGCCCGGGCGCGCATC 1943

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 17, 2003, 16:43:24 ; Search time 217.778 Seconds
(without alignments)
2280.750 Million cell updates/sec

Title: US-09-938-391-4

Perfect score: 966

Sequence: 1 HTHQDFQLVHLVALNSQP.....CRHAFVLCIENSVMTSFSK 184

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DRV=xlp
-Q/can2_1/USPTO_spool_p/US0938391/runat_04082003_130655_27077/app_query.fasta_1.718
-DB=N_Geneseq_19Jun03 -OFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -DOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0938391 @CCN 1.1 605 @runat_04082003_130655_27077 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-NO TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_19Jun03: *
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT: *
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT: *
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT: *
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT: *
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT: *
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT: *
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT: *
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT: *
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT: *
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT: *
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT: *
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT: *
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT: *
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT: *
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT: *
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT: *
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT: *
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT: *
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT: *
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: *
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT: *
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: *
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT: *
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: *
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	966	100.0	555	24	AAL46063	Canine endostatin
2	966	100.0	829	24	AAL46062	Canine pro-endosta
3	959	99.3	552	21	AAZ51309	Canine angiogenesi
4	835	86.4	546	21	AAZ27004	Human endostatin g
5	835	86.4	549	20	AAZ77719	Human endostatin D
6	835	86.4	549	21	AAA29884	Human angiogenesi
7	835	86.4	549	21	AAZ51291	Human angiogenesi
8	835	86.4	549	22	AAZ00867	Human gene fragmen
9	835	86.4	549	25	ABA00774	Human endostatin c
10	835	86.4	552	20	AAZ35375	SEQ ID 50 of W0991
11	835	86.4	552	21	AAZ62023	Nucleotide sequenc
12	835	86.4	552	21	AAA68203	Human endostatin c
13	835	86.4	552	21	AAZ50398	Human endostatin c
14	835	86.4	552	22	AAZ88289	Human endostatin c
15	835	86.4	552	24	ABK50685	cDNA encoding huma
16	835	86.4	555	22	AAI68529	Human vascular end
17	835	86.4	558	24	ABA99261	Human endostatin c
18	835	86.4	563	20	AAZ08750	Human endostatin c
19	835	86.4	816	21	AAZ64013	DNA encoding a hum
20	835	86.4	900	24	ABK09977	Synthetic plasmid
21	835	86.4	968	21	AAC62025	Nucleotide sequenc
22	835	86.4	1564	24	ABQ76740	DNA encoding human
23	835	86.4	3394	18	AAZ84484	Human alpha-1 coll
24	835	86.4	3394	24	AAZ78379	Human alpha (XVII
25	835	86.4	3394	24	ABN95680	Gene #2178 used to
26	835	86.4	4551	24	ABV94763	Human pancreatic c
27	835	86.4	4551	24	ABN85301	Human collagen XVI
28	835	86.4	4875	24	ABQ54955	Human ovarian anti
29	835	86.4	5408	20	AAZ77720	Human collagen 18
30	831	86.0	555	20	AAZ84635	Mouse endostatin c
31	831	86.0	555	20	AAZ79949	Endostatin coding
32	831	86.0	555	22	AAC88290	Murine endostatin
33	831	86.0	558	22	AAZ18701	Mouse endostatin e
34	831	86.0	558	24	ABQ76173	Rat endostatin enc
35	831	86.0	558	24	ABQ76178	Rat endostatin enc
36	831	86.0	565	20	AAZ58740	DNA encoding anti-
37	831	86.0	573	21	AAZ67777	Murine endostatin
38	831	86.0	573	24	ABK47719	cDNA encoding mous
39	830	85.9	641	22	AAZ79104	Human endostatin e
40	830	85.9	641	24	AAZ55454	Specific tumour ce
41	830	85.9	641	24	AAZ44000	Reproductive recom
42	830	85.9	641	24	ABQ76079	Anticancer gene-as
43	828	85.7	552	20	AAZ77715	Murine endostatin
44	828	85.7	552	21	AAZ51299	Murine angiogenesi
45	828	85.7	568	20	AAZ08749	Mouse endostatin c

ALIGNMENTS

RESULT 1
AAL46063
ID AAL46063 standard; cDNA; 555 BP.
XX
AC AAL46063;
XX
DT 19-JUL-2002 (first entry)
XX
DE Canine endostatin coding sequence.

Dog; pro-endostatin; endostatin; angiogenesis; cancer; gene therapy;
KW psoriasis; rheumatoid arthritis; retinopathy; macular degeneration;
KW corneal graft rejection; neovascular glaucoma; retrolental fibroplasia;
KW rubeosis; Osler-Weber Syndrome; myocardial angiogenesis;
KW plaque neovascularisation; telangiectasia; haemophilic joints;
KW angiodysplasia; wound granulation; coronary collateral;
KW cerebral collateral; arteriovenous malformation;

XX Claim 2; Fig 2; 56pp; English.
 PS The present invention provides the protein and coding sequences of canine
 CC pro-endostatin and endostatin. The sequences can be used in the treatment
 CC and diagnosis of angiogenesis related disorders, including cancer,
 CC rheumatoid arthritis, psoriasis, retinopathy, macular degeneration,
 CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasia,
 CC rubecosis, Osler-Webber Syndrome, myocardial angiogenesis, plaque
 CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,
 CC wound granulation, coronary collaterals, cerebral collaterals,
 CC arteriovenous malformations, ischemic limb angiogenesis, diabetic
 CC neovascularisation, and fractures. The present sequence is the canine
 CC pro-endostatin coding sequence.
 XX SQ Sequence 829 BP; 124 A; 314 C; 278 G; 113 T; 0 other;

Alignment Scores:
 Pred. No.: 5.47e-86 Length: 829
 Score: 965.00 Matches: 184
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

US-09-938-391-4 (1-184) x AAL46062 (1-829)

QY 1 HistHrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
 DB 139 CACACCCACGAGACTTCAGCTGGTGGTGGACCTGGTGGCCCTGAACAGCCGCGAGCG 198
 QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnAlaArgAlaAla 40
 DB 199 GCGGCGATCGAGGAGATCCGGGAGCGGACTTCAGTGGTTCAGCAGCGCGCGCGCG 258
 QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerile 60
 DB 259 GGGCTGGCGCGCACCTTCGGGCGCTTCCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 318
 QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
 DB 319 GTGCGCGCGCGAGCGAGCGCGCGGGTGGCGGCTGCTCAACCTCAGGAGCGAGTGTCTTC 378
 QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlyGlnLeuLysProGlyAlaArgile 100
 DB 379 CCCAGCTGGAGGCTTATCTCGGCTCGAGGCGCAGCTGAGCCCGGGCGCGCATC 438
 QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
 DB 439 TTCTCTTTCGCGGACAGATGTCTGTGACGACCCCGCTGGCCCGGAGAGCGGTGTGG 498
 QY 121 HisGlySerAspProSerGlyArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
 DB 499 CACGGTCCGACCCAGCGCGCGCGCGCTGACCGACAGCTACTGCGAGACGTGGCGGAGC 558
 QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyValGluLeuGluGln 160
 DB 559 GAGGCCCCCGCGGCGCGCGCGCGCGCGCTGCTGCTGCTGCGCGCGCGCGCGCGCGAG 618
 QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
 DB 619 GAGGCG 678
 QY 181 SerPheSerLys 184
 DB 679 TCCTTCTCCAAG 690

RESULT 3
 AA251309
 ID AA251309 standard; DNA; 552 BP.
 XX
 AC AA251309;
 XX

DT 06-JUN-2000 (first entry)
 XX Canine angiogenesis inhibitor, endostatin DNA.
 DE
 XX
 KW Canine; immunoglobulin Fc fragment; endostatin; immunofusin;
 KW angiogenesis; inhibitor; cytostatic; antirheumatoid; antiarthritic;
 KW antipsoriatic; antidiabetic; ophthalmological; immunosuppressant;
 KW vasotropic; vulnerary; treatment; antiarteriosclerosis; tumour;
 KW metastasis; atherosclerosis; psoriasis; rheumatoid arthritis;
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
 KW wound granulation; keloid scar; gene therapy; ds.
 XX
 OS Canis familiaris.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..552
 FT /*tag= a
 FT /product= "Endostatin"
 FT /note= "Does not include stop codon"
 FT /partial
 XX
 PN WO200011033-A2.
 XX
 PD 02-MAR-2000.
 XX
 PF 25-AUG-1999; 99WO-US19329.
 XX
 PR 25-AUG-1999; 98US-0097883.
 XX
 PA (LEXI-) LEXINGEN PHARM CORP.
 XX
 PI Lo K, Li Y, Gillies SD;
 XX
 DR WPI; 2000-237616/20.
 DR P-PSDB; AAY70265.
 XX
 PT Novel fusion protein of angiotensin or endostatin and an immunoglobulin
 PT Fc region, useful for treating conditions mediated by angiogenesis,
 PT such as rheumatoid arthritis, tumors and macular degeneration -
 XX
 PS Example 8; Pages 58-59; 68pp; English.
 XX
 CC The patent discloses a DNA molecule encoding a fusion protein comprising
 CC a signal sequence, an immunoglobulin Fc region, and an angiogenesis
 CC inhibitor selected from angiotensin, endostatin, a plasminogen fragment
 CC having angiotensin activity, a collagen XVIII fragment having endostatin
 CC activity, or combinations of them. The fusion protein (immunofusin) is
 CC used to inhibit angiogenesis and to treat diseases or conditions mediated
 CC by angiogenesis. Conditions that may be treated include solid tumours,
 CC blood born tumours, tumour metastasis, benign tumours including
 CC haemangiomas, acoustic neuromas, neurofibromas, trachomas and pyrogenic
 CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases
 CC e.g. diabetic retinopathy, retinopathy of prematurity, macular
 CC degeneration, corneal graft rejection, neovascular glaucoma, retrolental
 CC fibroplasia, rubecosis and Osler-Webber syndrome; myocardial angiogenesis,
 CC plaque neovascularisation, telangiectasia, haemophilic joints,
 CC angiofibroma, wound granulation, and excessive or abnormal stimulation of
 CC endothelial cells, intestinal cells, atherosclerosis, scleroderma and
 CC hypertrophic scars, i.e. keloid scars. The DNA constructs may be used
 CC in gene therapy. The present sequence is a DNA encoding canine
 CC endostatin used in the construction of immunofusin containing canine
 CC immunoglobulin Fc fragment.
 XX
 SQ Sequence 552 BP; 77 A; 204 C; 190 G; 81 T; 0 other;

Alignment Scores:

Pred. No.: 1.6e-85 Length: 552
 Score: 959.00 Matches: 183
 Percent Similarity: 99.46% Conservative: 0
 Best Local Similarity: 99.46% Mismatches: 1
 Query Match: 99.28% Indels: 0
 DB: 21 Gaps: 0


```
US-09-938-391-4 (1-184) x AA251309 (1-552)
QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
Db 1 CACACCCACAGGACTTCCAGCGCGTGTGCACCTGGTGGCCCTGAACAGCCGCGAGCG 60
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnAlaArgAla 40
Db 61 GCGCGCATGCGAGCATCCGGGAGCGGACTTCCAGTGTCTCAGCAGGCGCGCGCG 120
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIle 60
Db 121 GGGCTGGCGGCACCTTCCGGGCTTCTCTGCTGCGGCTGCAGGACCTCTACAGCATC 180
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
Db 181 GTGCGCGCGCGCCAGCCGCGCGGTGCGCGTGCACCTCAGGACGAGGTGCTCTTC 240
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 100
Db 241 CCAGCTGGAGGCTTATTTCTCGGCTCCGAGGCGAGTGAAGCCCGGGCGCGCATC 300
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 301 TTCTCTTCGACGCGCAGAGATGCTCTGCAGCACCCCGCTGGCCCGGAGAGCGGTG 360
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
Db 361 CACGGCTCCGACCCACCGCGCGCGCTTACCCGACAGCTACTTGCAGACGCTGGCGGAG 420
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
Db 421 GAGGCCCCGCGCGCGCCAGCGGCTGCTGCTGCTGCGGCGAGGCTGCTGGAGCAG 480
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
Db 481 GAGGCGCGGAGCTGCGCGCACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 181 SerPheSerLys 184
Db 541 TCCTTCTCCAAG 552
RESULT 4
AA227004
ID AAA27004 standard; DNA; 546 BP.
XX AAA27004;
XX AC
XX 11-AUG-2000 (first entry)
XX DE
XX Human endostatin gene.
XX Human; endothelial cell proliferation inhibitor; collagen XVIII;
XX angiogenesis inhibitor; anti-tumour; cytostatic; antiproliferative;
XX vasotrophic; dermatological; ophthalmological; vulnery;
XX antiarteriosclerotic; antidiabetic; haemostatic; contraceptive;
XX ocular angiogenic disease; atherosclerosis; scleroderma;
XX myocardial angiogenesis; telangiectasia; angiofibroma;
XX wound granulation, ds.
XX OS
XX Homo sapiens.
XX FH
XX Key Location/Qualifiers
XX CDS 1..546
XX FT /*tag= a
XX FT /product= "Endostatin protein"
XX FT
XX PN
XX WO200026368-A2.
XX XX
XX 11-MAY-2000.
XX XX
XX 01-NOV-1999; 99WO-US25605.
XX PF
```

```
XX 30-OCT-1998; 98US-0106343.
PR 20-MAY-1999; 99US-0315689.
XX (CHIL-) CHILDRENS MEDICAL CENT.
XX O'Reilly MS, Folkman MJ;
PI WPI; 2000-365617/31.
DR P-PSDB; AAY94323.
XX Novel endostatin capable of inhibiting endothelial cell proliferation
PT and angiogenesis, useful for treating angiogenesis-dependent cancers
PT and as birth control agents
XX Claim 10; Page 39; 68pp; English.
XX The present sequence encodes an endostatin protein which is the carboxy
CC terminal protein of human collagen XVIII. Recombinant mouse endostatin
CC (20 mg/kg) was administered subcutaneously to mice implanted with Lewis
CC lung carcinomas. There was tumour mass regression non-detectable levels
CC after 12 days of therapy due to the angiogenesis inhibitory activity of
CC endostatin. Thus the protein is useful for treatment of angiogenesis-
CC dependent cancers. The polynucleotide and polypeptide sequences of this
CC endostatin are useful for treating and diagnosis of tumours, ocular
CC angiogenic diseases, Osler-Webber syndrome, myocardial angiogenesis,
CC plaque neovascularisation, telangiectasia, haemophilic joints,
CC angiofibroma and wound granulation, for treatment of diseases related to
CC excessive or abnormal stimulation of endothelial cells e.g. intestinal
CC adhesions, atherosclerosis, scleroderma. The protein may also be useful
CC as a birth control agent by reducing or preventing uterine
CC vasularisation. The gene for endostatin may be isolated from cells or
CC tissue that express high levels of endostatin, eg. tumour cells, by
CC generating cDNA from mRNA using reverse transcriptase and then amplifying
CC the DNA sequence.
XX SQ Sequence 546 BP; 80 A; 196 C; 177 G; 93 T; 0 other;
Alignment Scores:
Pred. No.: 2,65e-73 Length: 546
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: 21 Gaps: 0
US-09-938-391-4 (1-184) x AAA27004 (1-546)
QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
Db 1 CACACCCACCGGACTTCCAGCGCGTGTCCACCTGTTGCTCCACAGCCCTGTCAC 60
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnAlaArgAla 40
Db 61 GCGCGCATGCGGGCATCCGGGGCCGACTTCCAGTGTCTTCCAGCAGCGCGCGCGT 120
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIle 60
Db 121 GGGCTGGCGGGCACCTTCCGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
Db 181 GTGCGCGCTGCGAGCGCGCGCGCGCTGCGCGCTCAGCTCAGGACGAGGTGCTGTT 240
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 100
Db 241 CCAGCTGGAGGCTCTGTCTTCTCAGGCTCTGAGGCTGAGGCTGAAGCCCGGGCGAGCATC 300
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 301 TTCTCTTTCGCGGCAAGAGCGTCTGAGGACCCCGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
Db 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
```



```
Db 361 CATGGCTCGACCCCAACGGCGCAGGCTGACCGAGAGCTACTGTGAGACGCTGGCGGACG 420
Qy 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGluGln 160
Db 421 GAGGCTCCCTCGGCCACGGGCCAGGCTCTCGTGTCTGGGGGAGGCTCTCTGGGGGAG 480
Qy 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
Db 481 AGTGGCGGAGCTGCCATCACGCTTACATCGTGTCTGCTGATTGAGAACAGCTTCATGACT 540
Qy 181 Ser 181
Db 541 GCC 543

RESULT 5
AAX77719
ID AAX77719 standard; DNA; 549 BP.
XX
AC AAX77719;
XX
DT 10-AUG-1999 (first entry)
XX
DE Human endostatin DNA coding region fragment.
XX
KW Plasminogen; human; angiotensin; endostatin; gene therapy; vector;
KW anti-angiogenic; attenuation; cytostatic; anti-diabetic; ophthalmology;
KW tumour growth; solid tumour; diabetic retinopathy; retina; ss.
OS Homo sapiens.
XX
PN WO926480-A1.
XX
PD 03-JUN-1999.
XX
PF 20-NOV-1998; 98WO-US24950.
XX
PR 20-NOV-1997; 97US-0975424.
XX
PA (GENE-) GENETIX PHARM INC.
PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX
PI Bachelot T, Leboulch P, Pawliuk RJ;
XX
DR WPI; 1999-357696/30.
DR P-PSDB; AAY08693.
XX
PT Anti-angiogenic gene therapy vectors
XX
PS Disclosure; Page 74; 83pp; English.
XX
CC This invention describes a novel viral gene therapy vector comprising a
CC nucleic acid molecule encoding an anti-angiogenic polypeptide chosen
CC from human or murine angiotensin, human or murine endostatin and
CC angiogenesis-inhibiting fusions and fragments, where the viral vector-is
CC sufficiently attenuated for use in human gene therapy. The products of
CC the invention have anti-angiogenic, cytostatic, anti-diabetic and
CC ophthalmological activity. The vector is used in gene therapy for
CC inhibiting tumour growth in humans harbouring a solid tumour. The vector
CC expresses an anti-angiogenic polypeptide. An additional use comprises
CC treatment of diabetic retinopathy, where the anti-angiogenic polypeptide
CC inhibits angiogenesis in the vicinity of the retina. The vector is
CC administered to cells ex vivo and then administered to the patient.
XX
SQ Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;

Alignment Scores:
Pred. No.: 2,67e-73 Length: 549
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: 20 Gaps: 0
```

```
US-09-938-391-4 (1-184) x AAX77719 (1-549)
Qy 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
Db 1 CACGCCACCGGAGCTTCACCGCGTGTCTCACCTGGTTCGCTCAACAGCCCCCTGTCA 60
Qy 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
Db 61 GCGCGCATCGCGGCGCATCCGCGGGCGGCGGCTTCACAGTGTCTCCAGCGCGGCGCGTG 120
Qy 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 121 GGGCTGGCGGGCACCTTCGCGCCTTCTGTCTCTCGCGCTGCAGGACCTGTACAGCATC 180
Qy 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
Db 181 GTGCGCGCTGGCGACCGCGGCGGCGGCTGCGCATCTCACTCAAGGACGAGTGTCTTT 240
Qy 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnGlnLeuLeuPheProGlyAlaAla 100
Db 241 CCCAGCTGGAGGCTCTGTCTCAGGCTCTGAGGCTCGGCTGAGCGCGGCGGCGGCGCATC 300
Qy 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 301 TTCTCCTTTGACGGCAAGGACGCTCTGAGGCAACCCCTGCGCCCGAGAGAGCGGTGG 360
Qy 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
Db 361 CATGCTCGAGCCCAACGCGGCGGAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGACG 420
Qy 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGluGln 160
Db 421 GAGGCTCCCTCGGCCACGGGCCAGGCTCTCTCGCTGGGGGCGGAGCTCTCTGGGGCAG 480
Qy 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
Db 481 AGTGGCGGAGCTGCCATCACGCTTACATCGTGTCTCTGCTTGCATTGAGAACAGCTTCATGACT 540
Qy 181 Ser 181
Db 541 GCC 543

RESULT 6
AAX29884
ID AAA29884 standard; cDNA; 549 BP.
XX
AC AAA29884;
XX
DT 22-AUG-2000 (first entry)
XX
DE Human angiogenesis inhibiting factor 1 encoding cDNA.
XX
KW Human; angiogenesis inhibiting factor 1; IAP-1; tumour; antibody;
KW abnormal vessel disease; ss.
OS Homo sapiens.
XX
PN CN1244536-A.
XX
PD 16-FEB-2000.
XX
PF 10-AUG-1998; 98CN-0117150.
XX
PR 10-AUG-1998; 98CN-0117150.
XX
PA (ONCO-) INST ONCOLOGY UNDER TUMOR HOSPITAL CHINE.
XX
PI Yang Z, Guo W;
XX
DR WPI; 2000-388168/34.
DR P-PSDB; AAY90771.
XX
```

PT Angiogenesis inhibiting factor 1 and its derivative useful for treating
tumors -
XX
PS Example 1; Fig 5; 41pp; Chinese.

XX The present sequence encodes an angiogenesis inhibiting factor (1),
CC designated IAF-1. The present invention also describes: (1) preparation
CC of (1) and its derivative; (2) an IAF binding acceptor and its
CC preparation; and (3) an IAF antibody. (1) is useful for preparing new
CC biological preparations for effectively treating various tumours and
CC abnormal-vessel diseases. The IAF antibody is preferably a polyclonal
CC antibody, mosaic antibody, single stranded antibody and human originated
CC antibody.

SQ Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;

Alignment Scores:
Pred. No.: 2,67e-73 Length: 549
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: 21 Gaps: 0

US-09-938-391-4 (1-184) x AAA29884 (1-549)

QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
DB 1 CACAGCCACCGCGACCTTCACGCGGTCCTCACCTGGTTCACAGCCCGCTGTCA 60
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaA 40
DB 61 GCGCGCATCGGGGACATCCCGGGGCGGACCTTCAGTGTTCACGAGCGCGGCGTG 120
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
DB 121 GGGCTGGCGGACCTTCCTCGCGCTTCCTGCTCGCGCTGCGAGACCTGTACAGCATC 180
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 181 GTCCGCGCTGCCACCGCGCGCGTCCGCTCAACCTCAAGGACGAGCTGTGTTT 240
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyAlaArgIle 100
DB 241 CCAGCTGGAGGAGCTGTCTCAGGCTCTGAGGTCCTGAGGCGCGGGGACGACATC 300
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
DB 301 TTCTCTTTGACGGCAAGGACGTCTTGAGGACCCACCTGGCGCCGAGAGCGGTGG 360
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
DB 361 CATGGCTCGGACCCCAACGCGCGCGAGCTGACGAGAGCTACTGTGAGACGTGGCGGACG 420
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGluGln 160
DB 421 GAGGCTCCCTCGGCGACCGGCGGAGGCTCTCTGCTGGGGGCGAGGCTCTGGGGCGAG 480
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuValLeuValLeuValMetThr 180
DB 481 AGTGGCGGAGCTGCCATCAGCGCTACATCGTCTGCTGCTGATTCAGAAACAGCTTCATGACT 540
QY 181 Ser 181
DB 541 GCC 543

RESULT 7
AAZ51291
ID AAZ51291 standard; cDNA; 549 BP.

XX AAZ51291;

DT 06-JUN-2000 (first entry)

XX Human angiogenesis inhibitor, endostatin cDNA.
DE
XX Human; immunoglobulin gamma Fc fragment; endostatin; immunofusin;
KW angiogenesis; inhibitor; cytostatic; antirheumatoid; antiarthritic;
KW antipsoriatic; antidiabetic; ophthalmological; immunosuppressant;
KW metastatic; vulnery; treatment; antiarteriosclerosis; tumour;
KW metastasis; atherosclerosis; psoriasis; rheumatoid arthritis;
KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
KW wound granulation; keloid scar; gene therapy; ss.
XX Homo sapiens.

WO200011033-A2.

02-MAR-2000.

25-AUG-1999; 99WO-US19329.

25-AUG-1998; 98US-0097883.

(LEXI-) LEXINGEN PHARM CORP.

Lo K, Li Y, Gillies SD;

WPI; 2000-237616/20.

P-PSDB; AAY70252.

Novel fusion protein of angiotensin or endostatin and an immunoglobulin
FC region, useful for treating conditions mediated by angiogenesis,
such as rheumatoid arthritis, tumors and macular degeneration -

Claim 12; Pages 40-41; 68pp; English.

The patent discloses a DNA molecule encoding a fusion protein comprising
a signal sequence, an immunoglobulin Fc region, and an angiogenesis
inhibitor selected from angiotensin, endostatin, a plasminogen fragment
having angiotensin activity, a collagen XVIII fragment having endostatin
activity, or combinations of them. The fusion protein (immunofusin) is
used to inhibit angiogenesis and to treat diseases or conditions mediated
by angiogenesis. Conditions that may be treated include solid tumours,
blood born tumours, tumour metastasis, benign tumours including
haemangiomas, acoustic neuromas, neurofibromas, trachomas and pyrogenic
granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases
e.g. diabetic retinopathy, retinopathy of prematurity, macular
degeneration, corneal graft rejection, neovascular glaucoma, retrolental
fibroplasia, rubiosis and Osler-Webber syndrome, myocardial angiogenesis,
plaque neovascularisation, telangiectasia, haemophilic joints,
angiobroma, wound granulation, and excessive or abnormal stimulation of
endothelial cells, intestinal cells, atherosclerosis, sclerodermal and
hypertrophic scars, i.e. keloid scars. The DNA constructs may be used
in gene therapy. The present sequence is a cDNA encoding human
endostatin used in the construction of immunofusin containing human
immunoglobulin gamma (IgG) Fc fragment.

Note: This sequence is stated in claim 12 as being amino acid sequence
of plasminogen fragment, however, the rest of the specification
refers to this sequence as being nucleotide sequence of human endostatin.

SQ Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;

Alignment Scores:

Pred. No.: 2,67e-73 Length: 549
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13

```
Query Match: 86.44% Indels: 0
DB: 21 Gaps: 0
US-09-938-391-4 (1-184) x AA251291 (1-549)

Qy 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAAsnSerProGlnPro 20
Db 1 CACAGCCACCGGACTTCCAGCGGTGTCTCCACCTGGTGTGGCTCAACAGCCCCCTGTCA 60
Qy 21 GlyGlyMetArgGlyLeuAlaAspPheGlnCysPheGlnGlnAlaArgAlaAla 40
Db 61 GCGGCGATGCGGGGATCCGCGGGCGGACTTCCAGTGTCTCCAGGAGCGGGCGCGTG 120
Qy 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 121 GGGCTGGCGGGCACCTTCCGCGCTTCTGTCTCCGCGCTGCAGGACCTGTACAGCATC 180
Qy 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
Db 181 GTGCGCGGTGCGGACCGCGGACCGTGGCCATCGTCAACCTCAAGGACGAGCTGTGTT 240
Qy 81 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLeuValProGlyAlaArgIle 100
Db 241 CCCAGCTGGAGGCTCTGTCTTCAGGCTCTGAGGCTCCGCTGAAGCCCGGGCAGCATC 300
Qy 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 301 TTCTCTTTGACGGCAAGGACGCTCTGAGGCACCCACCTGGCCCCCAGAGAGCGTGTGG 360
Qy 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
Db 361 CATGCTCGGACCCCAACGCGGCGAGGCTGACCGAGAGCTACTGTGACAGCTGGCGGAGC 420
Qy 141 GluAlaProAlaAlaThrClyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGln 160
Db 421 GAGGCTCCCTCGGCGACGCGGCGAGGCTCTCGCTGCTGGGGGCGAGGCTCTGGGGGAG 480
Qy 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAAsnSerValMetThr 180
Db 481 AGTGCGGCGAGTGTCATCACCGCTACATCGTGTCTGCTGATGAGACAGCTTCATGACT 540
Qy 181 Ser 181
Db 541 GCC 543

RESULT 8
AAS00867
ID AAS00867 standard; DNA; 549 BP.
AC AAS00867;
DT 04-JUL-2001 (first entry)
DE Human gene fragment encoding Endostatin(TM) protein.
KW Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;
KW blood borne tumour; leukaemia; tumour metastasis; benign tumour;
KW haemangioma; acoustic neuroma; neurofibroma; trachoma; rubecosis;
KW pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;
KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
KW retinopathy of prematurity; macular corneal graft rejection;
KW neovascular glaucoma; retrolental fibroplasia; Osler-Webber Syndrome;
KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
KW haemophilic joint; angiofibroma; wound granulation; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT 1..540
FT /*tag= a
FT /*product= "Endostatin(TM) C-terminus minus 3"
FT /*partial
FT /*note= "Variant produced during fermentation
FT Alignment Scores: 2.67e-73 Length: 549
Pred. No.:
```

Score: 839.00 Matches: 154
Percent Similarity: 92.82% Conservatives: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: 22 Gaps: 0

US-09-938-391-4 (1-184) x AAS00867 (1-549)

```
QY 1 HisThrHisGlnAspPheGlnLeuValHisLeuValAlaLeuAsnSerProGlnPro 20
DB 1 CACAGCCACCGGACTTCAGCGCGTCTCCACCTGGTGGCTCAACAGCCCCCTGTCA 60

QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnAlaAla 40
DB 61 GCGGCGATGCGGGGATCCGCGGGGCGGACTTCCAGTGTCTCCAGCAGCGCGCGGTG 120

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSer 60
DB 121 GGGCTGGCGGGACCTTCGCGGCTTCTGTCTCGCGCTCGAGACCTGTACAGCATC 180

QY 61 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 181 GTGCGCGGTGCGAGCCGCGCGGCTGCTCACTGTCACCTCAGAGCAGTGTCTTT 240

QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLeuPheProGlyAlaArg 100
DB 241 CCAGCTGGAGGCTCTGTCTCAGGCTCTGAGGCTCGCTGAAGCCGCGGCGACGATC 300

QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerVal 120
DB 301 TTCTCTTTGACGGCAAGACGCTCTGAGGACCCGCTGCTGCTGGCGGCGAGCGTGTG 360

QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGlnThrTrpArg 140
DB 361 CATGGCTCGAGCCCAACCGCGCGAGCTGACCGAGAGCTACTGTGACGCTGGCGGAC 420

QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGlu 160
DB 421 GAGGCTCCCTCGGACACGGCGCAGGCTCTCTGCTGCTGGGGGCGAGGCTCTGGGG 480

QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMet 180
DB 481 AGTGGCGGAGCTGCCATCAGCGCTACATCGTCTGTGCTGTGATTCGAGAACAGTTC 540

QY 181 Ser 181
DB 541 GCC 543
```

RESULT 9

ABA00774

ID ABA00774 standard; cDNA; 549 BP.

XX AC ABA00774;

XX XX 18-MAR-2003 (first entry)

XX DE Human endostatin coding sequence.

XX KW Gene; human; plasminogen; angiostatin; neovascularisation;

XX KW kringle domain; cell proliferation; viral vector;

XX KW replication-defective; cancer; tumour; ss.

XX OS Homo sapiens.

XX XX Key Location/Qualifiers

FT CDS 1..549

FT /tag= a

FT /product= "Endostatin"

XX PN WO20028173-A2.

XX PD 07-NOV-2002.

XX XX

PF 29-APR-2002; 2002WO-US13461.

XX 30-APR-2001; 2001US-287673P.

PR 05-APR-2002; 2002US-370634P.

XX (CELL-) CELL GENESYS INC.

PA Chang B, Wu WW, Macarthur J, Patel S, Jooss K, Mendez M;

XX MPI; 2003-129131/12.

XX P-FSDB; AAG79753.

XX New recombinant viral vector expressing human angiostatin useful for

XX inhibiting angiogenesis in a mammalian subject with cancer or tumor

XX Example 4; Page 80-82; 83pp; English.

XX This sequence encodes endostatin. Endostatin is a 20 kD C-terminal

XX fragment of collagen XVIII that inhibits angiogenesis. The endostatin

XX coding sequence may be used in the recombinant viral vector of the

XX invention for obtaining angiostatin activity. The vector comprises a

XX promoter capable of expressing human angiostatin operably linked to a

XX structural gene encoding one or more domains of human angiostatin.

XX The vector, which may be a replication-defective viral vector, is useful

XX for inhibiting angiogenesis in a mammal, especially with cancer or a

XX tumour.

XX Note: This sequence is given incorrectly in the sequence listing of

XX the specification as an amino acid sequence.

XX SQ Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;

Alignment Scores:

Pred. No.: 2,67e-73 Length: 549
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservatives: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: 25 Gaps: 0

US-09-938-391-4 (1-184) x ABA00774 (1-549)

QY 1 HisThrHisGlnAspPheGlnLeuValHisLeuValAlaLeuAsnSerProGlnPro 20

DB 1 CACAGCCACCGGACTTCAGCGGTGCTCCACCTGGTGGCTCAACAGCCCCCTGTCA 60

QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnAlaAla 40

DB 61 GCGGCGATGCGGGGATCCGCGGGGCGGACTTCCAGTGTCTCCAGCAGCGCGCGGTG 120

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSer 60

DB 121 GGGCTGGCGGGACCTTCGCGGCTTCTGTCTCGCGCTCGAGACCTGTACAGCATC 180

QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80

DB 181 GTGCGCGTGGCGGACCGCTGCGGCGAGCTGCTCAACCTCAAGCAGAGTGTCTGT 240

QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLeuPheProGlyAlaArg 100

DB 241 CCAGCTGGAGGCTCTGTCTCAGGCTGTGAGGCTCGCTGAAGCCCGGGCAGCATC 300

QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerVal 120

DB 301 TTCTCTTTGACGGCAAGACGCTCTGAGGACCCGCTGCTGCTGGCGGCGAGCGTGTG 360

QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGlnThrTrpArg 140

DB 361 CATGGCTCGAGCCCAACCGCGCGAGCTGACCGAGAGCTACTGTGAGACGCTGGCGG 420

QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGlu 160

DB 421 GAGGCTCCCTCGGCGGCGGCGGCTCTCTGCTGCTGGGGGCGAGGCTCTCTGGGG 480

QY 161 GluAlaLaSerCysArgHisAlaPheValValLeuValCysIleGluAsnSerValMetThr 180
 Db 481 AGTGGCGGAGTGCCATCAGCCTACATCGTCTGCTGATTGAGAACAGCTTCATGACT 540
 QY 181 Ser 181
 Db 541 GCC 543

RESULT 10

AA35375

ID AAX35375 standard; DNA; 552 BP.

XX

AC AAX35375;

DT 16-JUL-1999 (first entry)

XX

DE SEQ ID 50 of WO9916889.

XX

KW Angiostatin; endostatin; interferon; thrombospondin;

KW interferon-inducible protein; platelet factor 4; anti-angiogenic;

KW anti-tumor; multifunctional protein; angiogenic-mediated disease;

KW cancer; diabetic retinopathy; macular degeneration; arthritis;

KW tumor cell production; ss.

XX

OS Homo sapiens.

XX

PN WO9916889-A1.

XX

PD 08-APR-1999.

XX

PF 30-SEP-1998; 98WO-US20464.

XX

PR 01-OCT-1997; 97US-0060609.

XX

PA (SEAR) SEARLE & CO G D.

XX

PI Bolanowski MA, Caparon MH, Casperson GF, Gregory SA;

PI Klein BK, McKearn JP;

XX

DR WPI; 1999-255098/21.

XX

PT New multifunctional proteins useful for treating angiogenic-mediated

PT diseases

PS Disclosure; Page 85; 121pp; English.

XX

CC The specification describes multifunctional proteins which comprise combinations of angiostatin, endostatin, interferon, thrombospondin, interferon-inducible protein and platelet factor 4, and have

CC anti-angiogenic and/or anti-tumor activity. The multifunctional protein

CC may exhibit useful properties such as having similar or greater

CC biological activity when compared to a single factor or by having

CC improved half-life or decreased adverse side effects, or a combination

CC of these properties. The proteins can be used for treating an

CC angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular

CC degeneration, or arthritis. They can also be used for inhibiting the

CC production of tumor cells (characteristic of lung, breast, ovarian,

CC prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma,

CC hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor

CC growth. The present sequence is used in the course of the invention.

XX

SQ Sequence 552 BP; 83 A; 196 C; 179 G; 94 T; 0 other;

Alignment Scores:

Pred. No.: 2,69e-73 Length: 552
 Score: 835.00 Matches: 154
 Percent Similarity: 92.82% Conservative: 14
 Best Local Similarity: 85.08% Mismatches: 13
 Query Match: 86.44% Indels: 0
 DB: 20 Gaps: 0

US-09-938-391-4 (1-184) x AAX35375 (1-552)

QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
 Db 1 CACGCCACCGCGACTTCAGCGGTGCTCCACCTGGTTGGCTCAACAGCCCCCTGTCA 60
 QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
 Db 61 GCGCGCATGCGGGGCATCCGCGGGCGGACTTCCAGTGTCTCCAGCAGCGCGGCGGTG 120
 QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIle 60
 Db 121 GGGCTGGCGGGCACCCTTCGCGGCTTCTGCTCGCGCTGCTGAGACCTGTACAGCATC 180
 QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
 Db 181 GTGCGCGTGCAGACCGCGCAGCGTGCCTCAACCTCAGGACGAGCTGTGTGT 240
 QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlnGlyGlnLeuValProGlyAlaArgIle 100
 Db 241 CCCAGCTGGAGGCTCTGTTCTCAGGCTCTGAGGCTCGCTGAAGCCCGGGGCACGCATC 300
 QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
 Db 301 TTCTCTTTGACGGCAAGACGCTCTGAGGACCCCACTGGCCCCAGAGAGCGTGTGG 360
 QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
 Db 361 CATGGCTCGGACCCCAACGCGCGCAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGACG 420
 QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGln 160
 Db 421 GAGGCTCCCTCGGCCACCGCGCAGGCTCTGCTGCTGGGGGGCAGGCTCTCGGGGCGAG 480
 QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
 Db 481 AGTGGCGGAGCTGCCATCAGCCCTACATGCTGCTCTGCAITGAGAACAGCTTCATGACT 540
 QY 181 Ser 181
 Db 541 GCC 543

RESULT 11

AA362023

ID AAC62023 standard; DNA; 552 BP.

XX

AC AAC62023;

XX

DT 06-MAR-2001 (first entry)

XX

DE Nucleotide sequence of human endostatin encoded by plasmid pMALCH#15.

XX

KW Streptomyces sp. strain C5; SupA; S. venezuelae; alpha-amylase;

KW endostatin; cancer; tumour growth; angiogenesis; ss.

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..552

FT /tag= a

FT /product= "endostatin"

XX

PN WO200060945-A1.

XX

PD 19-OCT-2000.

XX

PF 12-APR-2000; 2000WO-US09747.

XX

PR 13-APR-1999; 99US-0129084.

XX

PA (MERI) MERCK & CO INC.

XX

PI Desanti CL, Strohl WR;

XX

DR WPI; 2000-686970/67.

DR P-PSDB; AAB30493.
XX
PT Preparation of soluble recombinant endostatin involves transforming
PT Streptomyces host with expression vector comprising nucleotide
PT sequence encoding endostatin operably linked to linker and leader
PT peptide -
XX
PS
XX Example 1; Fig 6; 57pp; English.
XX
CC The present sequence encodes human endostatin. The protein is expressed
CC in Streptomyces. Leader sequences of Streptomyces sp. strain C5
CC SnpA and S. venezuelae alpha-amylase proteins are linked to the
CC N-terminal of endostatin. This ensures that endostatin protein is
CC produced as a secreted, soluble protein which needs no refolding, is
CC stable in the fermentation broth and is produced in large quantities.
CC The method is used for preparing soluble recombinant human, murine or
CC primate endostatin, which is useful in the treatment of cancer,
CC inhibition of tumour growth, inhibition of angiogenesis, isolation of
CC receptors for endostatin and for identification of anti-angiogenic
CC compounds in assays. The endostatin protein is produced as a secreted,
CC soluble protein which needs no refolding, is stable in the fermentation
CC broth and is produced in large quantities. Streptomyces are amenable
CC for cultivation in large fermentations allowing for large quantities of
CC soluble endostatin to be produced.
XX
SQ Sequence 552 BP; 83 A; 196 C; 179 G; 94 T; 0 other;

Alignment Scores:
Pred. No.: 2,69e-73 Length: 552
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: 21 Gaps: 0

US-09-938-391-4 (1-184) x AAC62023 (1-552)

QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
DB 1 CACAGCCACCGGACTTCCAGCGGTGCTCCACCTGTTGGCTCAACAGCCCCCTGTCA 60
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
DB 61 GCGCGCATCGGGGACCTCCGCGGGCGGACTTCCAGTGTCTCCAGCAGCGCGGGCGGTG 120
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTrpSerile 60
DB 121 GGGCTGGCGGGCACCTTCCGCGGCTTCTGCTCTCGCGCTGCGAGACCTGTACAGCATC 180
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 181 GTGGCGCGTGGCGACCGCGCAGCGTGGCCATCGTCAACTCAAGACAGAGCTGTGTT 240
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLeuPheProGlyAlaArgile 100
DB 241 CCAGCTGGAGGCTCTGTCTCAGGCTCTGAGGTCCTGAGGTCCTGAAGCCCGGGGACGATC 300
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArglySerValTrp 120
DB 301 TTCTCTTTGAGCGGAAGACGCTCTGAGGACCCACCTGGCCCCAGAGAGCGTGTGG 360
QY 121 HisGlySerAspProSerGlyArgLeuThrAspSerTyCysGluThrTrpArgThr 140
DB 361 CATGGCTCGAGCCCAACCGGGCGAGCTGACGAGAGTACTGTGAGACGTGGGGAGC 420
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGluGln 160
DB 421 GAGGCTCCCTCGGCCACCGGCGGAGGCTCTCGCTGCTGGGGGCGAGGCTCTGGGGCAG 480
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuValLeuValLeuValMetThr 180
DB 481 AGTGCCCGGAGGTGCCATCACGCTACATCGTGTGCTGTGATTGAGAACAGCTTCATGACT 540

QY 181 Ser 181
DB 541 GGC 543
RESULT 12
AAA68203
ID AAA68203 standard; cDNA; 552 BP.
XX
AC AAA68203;
XX 27-OCT-2000 (first entry)
XX Human endostatin cDNA sequence.
XX Angiogenesis-inhibiting protein receptor; angiogenesis; angiostatin;
KW endostatin; plasminogen; laminin; treatment; wound healing; solid tumour;
KW psoriasis; scleroderma; myocardial angiogenesis; Crohn's disease;
KW cerebral collateral; arteriovenous malformation; rubeosis; cancer;
KW diabetic retinopathy; arthritis; wound healing; peptic ulcer;
KW Helicobacter related disease; fracture; cat scratch fever; ss.
XX Homo sapiens.
XX
XX WO200032631-A2.
XX 08-JUN-2000.
XX 06-DEC-1999; 99WO-US28897.
XX 04-DEC-1999; 98US-0206059.
XX (ENTR-) ENTREMED INC.
XX MacDonald NJ, Sam KL;
XX WPI; 2000-412290/35.
XX New angiogenesis-inhibiting protein receptors, useful in methods for
PT treating diseases and processes that are mediated by angiogenesis, such
PT as solid tumours, psoriasis, scleroderma and myocardial angiogenesis -
XX Disclosure; Page 86; 100pp; English.
XX This invention relates to angiogenesis-inhibiting protein receptors, and
CC the DNA sequences encoding them. Angiogenesis is the generation of new
CC blood vessels into a tissue, and normally occurs in wound healing,
CC foetal and embryonal development and the formation of the corpus luteum,
CC endometrium and placenta. Angiostatin is a protein (see AAB16450 and
CC AAA68202) involved in angiogenesis, and has an amino acid sequence
CC similar to that of a plasminogen fragment (see murine plasminogen
CC AAB16490). Angiostatin has the ability to inhibit angiogenesis.
CC Endostatin is also an angiogenesis inhibiting protein (see AAB16451 and
CC AAA68203). Sequences AAA68242 and AAB16522 represent coding and protein
CC sequences of human laminin. Laminin is an angiostatin binding protein,
CC and some of the peptides of the invention share homology with regions of
CC laminin. Peptides AAB16452-B16521 (excluding AAB16490) are the
CC angiogenesis-inhibiting protein receptor fragments of the invention. The
CC peptides bind either angiostatin or endostatin and can be used in methods
CC for treating diseases and processes that are mediated by angiogenesis,
CC such as solid tumours, psoriasis, scleroderma, myocardial angiogenesis,
CC Crohn's disease, cerebral collaterals, arteriovenous malformations,
CC rubeosis, diabetic retinopathy, arthritis, wound healing, peptic ulcers,
CC Helicobacter related diseases, fractures, placenta and cat scratch
CC fever. They are useful for the detection and prognosis of cancer. DNA
CC sequences A628204-A628241 encode the peptides of the invention.
SQ Sequence 552 BP; 83 A; 196 C; 179 G; 94 T; 0 other;
Alignment Scores:
Pred. No.: 2,69e-73 Length: 552
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13

```
Query Match: 86.44% Indels: 0
DB: 21 Gaps: 0
US-09-938-391-4 (1-184) x AAA68203 (1-552)
QY 1 HisThrHisGlnAspPheGlnLeuValHisLeuValAlaLeuAsnSerProGlnPro 20
DB 1 CACAGCCACCGGACTTCCAGCGGTGTCTCCACCTGGTTGCGCTCAACAGGCCCTGTGCA 60
QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAla 40
DB 61 GCGGCATCGGGGATCCGCGGCGGACTTCCAGTGTCTCCAGAGCGGGCGCGTG 120
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
DB 121 GGGCTGGCGGGCACCTTCGCGCCCTTCTGTCTCTCGCGCTCAGGACCTGTACAGCATC 180
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 181 GTGCGCGGTGCGAGCGCGCGACCGTGGCCATCGTCAACCTCAAGGACGAGCTGTGTTT 240
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyAlaArgIle 100
DB 241 CCCAGCTGGAGGCTCTGTCTCAGGCTCTGAGGCTCGCGTGAAGCCCGGCGACGATC 300
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
DB 301 TTCTCTTTGAGCGGCAAGGAGCTCTCTGAGGCACCCACCTGGCCCCAGAGAGCTGTGG 360
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
DB 361 CATGCTCGGACCCCAACCGGGCGAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGAGC 420
QY 141 GluAlaProAlaAlaThrClyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
DB 421 GAGGCTCCCTCGGCCACCGGCGGAGGCTCTCGCTGCTGGGGGCGAGGCTCCCTGGGGG 480
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
DB 481 AGTGGCGGAGTGCGATCAGCCCTACATCGTGTCTCTGATTTGAGAACAGCTTCATGACT 540
QY 181 Ser 181
DB 541 GCC 543
RESULT 13
AAZ50398
ID AAZ50398 standard; cDNA; 552 BP.
XX AC AAZ50398;
XX XX Human endostatin cDNA.
XX DT 18-MAY-2000 (first entry)
XX DE
XX KW Human endostatin; angiogenesis; expression plasmid; tumour activity;
XX KW cancer; anti-angiogenic agent; solid tumour; lung metastatic tumour;
XX KW cytostatic; gene therapy; ss.
XX XX Homo sapiens.
XX OS
XX PN WO200006759-A2.
XX XX
XX PD 10-FEB-2000.
XX XX
XX PF 20-JUL-1999; 99WO-US16388.
XX XX
XX PR 27-JUL-1998; 98US-0094375.
XX XX
XX PA (VALE-) VALENTIS INC.
XX XX
XX PI Min W, Szymanski P, Mehrens D, Ralston R, Sullivan S;
XX ID AAC88289 standard; DNA; 552 BP.
```

```
DR WPI; 2000-183133/16.
XX Plasmids comprising tissue specific transcription elements linked to an
PT anti-angiogenic gene is useful transfection of cells and treatment of,
PT e.g. cancer
XX
XX Disclosure; Page 40; 103pp; English.
XX
CC The present sequence encodes human endostatin, an angiogenesis
CC inhibitor derived from human liver cDNA. This is the C-terminal
CC proteolytic fragment of collagen 18a. Endostatin can inhibit endothelial
CC cell proliferation in vitro and angiogenesis in vivo. This is used in
CC the construction of the expression plasmid for the treatment of
CC mammalian diseases, especially cancer. The plasmids can be used for
CC (in vivo) transfection of a cell in situ in order to modulate tumour
CC activity. Anti-angiogenic gene inhibits growth of solid tumour and lung
CC metastatic tumours by intravenous or intramuscular delivery.
XX
SQ Sequence 552 BP; 87 A; 223 C; 173 G; 69 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 2,69e-73 Length: 552
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: 21 Gaps: 0
```

US-09-938-391-4 (1-184) x AAZ50398 (1-552)

```
QY 1 HisThrHisGlnAspPheGlnLeuValHisLeuValAlaLeuAsnSerProGlnPro 20
DB 1 CACAGCCACCGGACTTCCAGCGGTGTCTCCACCTGGTTGCGCTCAACAGGCCCTGTGAGC 60
QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAla 40
DB 61 GCGGCATCGGGGATCCGCGGCGGACTTCCAGTGTCTCCAGAGCGGGCGCGTG 120
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
DB 121 GGGCTGGCGGGCACCTTCGCGCCCTTCTGTCTCTCGCGCTCAGGACCTGTACAGCATC 180
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 181 GTGCGCGGTGCGAGCGCGCGACCGTGGCCATCGTCAACCTCAAGGACGAGCTGTGTTT 240
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyAlaArgIle 100
DB 241 CCCAGCTGGAGGCTCTGTCTCAGGCGGAGCGGCGGCGGCTGGAAGCCCGGCGGCGATC 300
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
DB 301 TTCTCTTTGAGCGGCAAGGAGCTGTCTGAGGCACCCACCTGGCCCCAGAGAGCTGTGG 360
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
DB 361 CACGCGACGAGCCCAACCGGCGGCGCTGACCGAGAGCTACTGCGAGACCTGGCGGAC 420
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
DB 421 GAGGCGGCGGCGGCGGCGGCGGAGGAGGCTGTGGGGCGGCGGCTGTGGGGCGGAG 480
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
DB 481 AGCGCGGCGGAGTGCGATCAGCCCTACATCGTGTCTGTGATTCGAGAACAGCTTCATGACC 540
QY 181 Ser 181
DB 541 GCC 543
RESULT 14
AAC88289
ID AAC88289 standard; DNA; 552 BP.
```

```
XX AAC88289;
XX 02-MAR-2001 (first entry)
XX Human endostatin coding sequence SEQ ID NO: 1.
XX DE
XX Endostatin; antiangiogenic; angiogenesis; human; mouse; chicken;
XX KW cancer; inflammation; angiogenesis-dependent disease; ds.
XX OS Homo sapiens.
XX PN WO200067771-A1.
XX PD 16-NOV-2000.
XX PF 02-MAY-2000; 2000WO-US12063.
XX PR 06-MAY-1999; 99US-0132907.
XX PR 14-JUL-1999; 99US-0353333.
XX PA (BURN-) BURNHAM INST.
XX PI Vuori K;
XX DR WPI: 2001-040937/05.
XX DR P-PSDB; AAB49379.
XX Endostatin peptide comprising at least four endostatin amino acid
PT residues are e.g. angiogenesis inhibitors for treating cancer and
PT diabetic retinopathy -
PT PS Disclosure; Page 111-112; 146pp; English.
XX CC The present invention provides endostatin peptides which can be used in
CC the modulation of angiogenesis. This is useful in the treatment of
CC cancers, inflammation, rheumatoid arthritis, chronic articular
CC rheumatism, psoriasis, disorders associated with inopportune invasion of
CC vessels such as diabetic retinopathy, neovascular glaucoma, retinopathy
CC of prematurity, macular degeneration, corneal graft rejection,
CC retrolental fibroplasia, rubeosis, capillary proliferation in
CC atherosclerotic plaques and osteoporosis. Other angiogenesis-dependent
CC diseases include Osler-Webber syndrome, myocardial angiogenesis, plaque
CC neovascularisation, telangiectasia, haemophilic joints and wound
CC granulation. In addition, the peptides can be used as birth control
CC agents.
XX SQ Sequence 552 BP; 83 A; 196 C; 179 G; 94 T; 0 other;

Alignment Scores:
Pred. No.: 2,69e-73 Length: 552
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: 22 Gaps: 0

US-09-938-391-4 (1-184) x AAC88289 (1-552)
QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
DB 1 CACAGCCACCGGACCTCCAGCGGTGCTCCACCTGTTGGCTCAACAGCCCTGTCA 60
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
DB 61 GCGGCAATCGGGGACATCCGCGGGCGGACTTCCAGTGTTCAGCAGCGCGGGCGT 120
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTrSerile 60
DB 121 GGGCTGGCGGGACCTTCGGCGCTTCCTCGCGCTTCAGGACCTGTACAGCATC 180
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
|||||
```

```
DB 181 GTGCGCGTGGCCGACCGCGCGAGCCGTCGTCCTCAAGTCAAGGACGAGCTGTGT 240
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlnGlyGlnLeuValProGlyAlaArgIle 100
DB 241 CCCAGCTGGAGGCTCTGTCTCAGGCTCTGAGGCTCCGCTGAAGCCCGGCGCAGCATC 300
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
DB 301 TTCTCTTTGACGGCAAGGAGCTCTCTGAGGCACCCACCTGGCCCCCAGAGAGCGTGTGG 360
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
DB 361 CATGGCTCGGACCCCAACGGCGGAGCTCTGAGGCACCCACCTGGCCCCCAGAGAGCGTGTGG 420
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGln 160
DB 421 GAGGCTCCCTCGGCCACGGCCAGGCTCTCTCGCTGGGGGGGAGGCTCTCTGGGGCAG 480
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
DB 481 AGTGGCGGAGCTGCATCAGCCTACATCGTGTCTTGCATTGAGACAGCTTCATGACT 540
QY 181 Ser 181
DB 541 GCC 543
RESULT 15
ABK50685
ID ABK50685 standard; cDNA; 552 BP.
XX AC ABK50685;
XX DT 13-AUG-2002 (first entry)
XX DE cDNA encoding human endostatin.
XX KW Human; angiogenesis; PITSRE protein kinase; cancer; arthritis;
KW macular degeneration; diabetic retinopathy; angiogenic-related disease;
KW haemangioma; blood borne tumour; leukaemia; neovascularisation;
KW coronary collateral; cerebral collateral; neovascular glaucoma;
KW corneal disease; wound healing; Helicobacter related disease; fracture;
KW keloid; ovulation; menstruation; gene; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 1..552
XX FT /*tag= a
XX FT /partial
XX FT /product= "Endostatin"
XX FT /note= "This sequence lacks a start codon"
XX WO200230982-A2.
XX 18-APR-2002.
XX 15-OCT-2001; 2001WO-US32437.
XX 13-OCT-2000; 2000US-240127P.
XX (ENTR-) ENTREMED INC.
XX PI Sim KL, Liang H;
XX WPI; 2002-435440/46.
XX DR P-PSDB; AAU97132.
XX Regulating angiogenesis for treating scleroderma, leukaemia, keloids by
PT administering a protein that is homologous to PITSRE protein kinase
PT and an angiogenic factor or a protein kinase and its active fragments -
XX Disclosure; Fig 2B; 45pp; English.
XX
```


Search completed: August 17, 2003, 20:06:34
Job time : 220.778 secs

CC The present invention relates to methods and compositions of inhibiting
CC angiogenesis. The method comprises administering to a human or animal
CC a composition comprising a protein that is homologous to PITSRE
CC protein kinases (PK) and an angiogenic factor. The method is useful
CC for regulating angiogenesis related to cancer, arthritis, macular
CC degeneration, and diabetic retinopathy. The compositions are useful
CC for inhibiting angiogenic-related diseases. The method and
CC compositions are useful in treating diseases and processes that are
CC mediated by angiogenesis including haemangioma, solid tumours, blood
CC borne tumours, leukaemia, metastasis, telangiectasia, psoriasis,
CC scleroderma, pyogenic granuloma, myocardial angiogenesis, Crohn's
CC disease, plaque neovascularisation, coronary collaterals, cerebral
CC collaterals, arteriovenous malformations, ischaemic limb angiogenesis,
CC corneal diseases, rubecosis, neovascular glaucoma, diabetic retinopathy,
CC retrolental fibroplasia, arthritis, diabetic neovascularisation,
CC macular degeneration, wound healing, peptic ulcer, Helicobacter related
CC diseases, fractures, keloids, vasculogenesis, haematopoiesis, ovulation,
CC menstruation, placentalion, and cat scratch fever. The method of
CC the invention provides a therapy for cancer that has minimal side
CC effects. The present sequence encodes human endostatin which is
CC used to generate angiogenesis-inhibiting peptides.

XX
SQ Sequence 552 BP; 83 A; 196 C; 179 G; 94 T; 0 other;

Alignment Scores:

Pred. No.:	2,69e-73	Length:	552
Score:	835.00	Matches:	154
Percent Similarity:	92.82%	Conservative:	14
Best Local Similarity:	85.08%	Mismatches:	13
Query Match:	86.44%	Indels:	0
DB:	24	Gaps:	0

US-09-938-391-4 (1-184) x ABK50685 (1-552)

QY	1	HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro	20
DB	1	CACAGCCACCGGACCTTCAGCGGTGTCTCCACCTGGTGGCTCAACAGGCCCTGTGCA	60
QY	21	GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaAla	40
DB	61	GGCGCATGCGGGGCATTCGCGGGGCCGACTTCCAGTGCTTCCAGCGCGGGGCCGTG	120
QY	41	GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle	60
DB	121	GGGCTGGCGGGCACCTTCGCGGCTTCTGTCTCTCGCGCTGCAGACCTGTACAGCATC	180
QY	61	ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe	80
DB	181	GTGCGCGGTGCG	240
QY	81	ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyValaArgIle	100
DB	241	CCAGCTGGAGGCTCTGTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGGCACGCATC	300
QY	101	PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp	120
DB	301	TTCTCTTTGACGGCAGGACCTCTGAGGACCCACCTGCGCCCGAGAGAGCGTGTGG	360
QY	121	HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr	140
DB	361	CATGGCTCGGACCCCAACCGGGCGGAGGCTGACCGAGAGCTACTGTGACACGTGGCGG	420
QY	141	GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln	160
DB	421	GAGGCTCCCTCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG	480
QY	161	GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr	180
DB	481	AGTGGCGGAGCTGCCATCACGCTTACATCTGCTGCTTGCATTGAGAACAGCTTCA	540
QY	181	Ser 181	
DB		:::	
DB	541	GCC 543	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 17, 2003, 19:49:34 ; Search time 57.7778 Seconds
(without alignments)
1405.636 Million cell updates/sec

Title: US-09-938-391-4
Perfect score: 966
Sequence: 1 HTHQDFQLVHLVALNSQP.....CRHAFVLCIENSVMTSFSK 184

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -MODEL=frame.p2n.model -DEV=xl

-Q=/cgn2_1/USFTO.spool_p/US0938391/runat_04082003_130656_27125/app.query.fasta_1.718
-DB=Issued Patents NA -OFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0938391@cgn.1.133 @runat_04082003_130656_27125 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA :
1: /cgn2_6/prodata/2/ina/5A COMB.seq :
2: /cgn2_6/prodata/2/ina/5B COMB.seq :
3: /cgn2_6/prodata/2/ina/6A COMB.seq :
4: /cgn2_6/prodata/2/ina/6B COMB.seq :
5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq :
6: /cgn2_6/prodata/2/ina/backfiles1.seq :

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	835	86.4	546	US-09-315-689-4	Sequence 4, Appli
2	835	86.4	552	US-09-206-059-30	Sequence 30, Appl
3	835	86.4	3394	US-08-159-784-4	Sequence 4, Appli
4	831	86.0	558	US-09-449-293-3	Sequence 3, Appli
5	831	86.0	558	US-09-775-325-3	Sequence 3, Appli
6	831	86.0	565	US-08-985-526-37	Sequence 37, Appl
7	831	86.0	573	US-09-561-500-12	Sequence 12, Appl
8	831	86.0	573	US-09-561-108-12	Sequence 12, Appl
9	831	86.0	573	US-09-561-528-12	Sequence 12, Appl
10	831	86.0	573	US-09-561-499-12	Sequence 12, Appl
11	828	85.7	4031	US-08-159-784-1	Sequence 1, Appli
12	817	84.6	534	US-09-315-689-6	Sequence 6, Appli

13	90.5	9.4	591	4	US-09-252-991A-15455	Sequence 15455, A
c 14	90.5	9.4	2319	4	US-09-252-991A-15284	Sequence 15284, A
c 15	90.5	9.4	2427	4	US-09-252-991A-15396	Sequence 15396, A
c 16	90	9.3	1413	4	US-09-252-991A-9741	Sequence 9741, Ap
c 17	90	9.3	2082	4	US-09-252-991A-9500	Sequence 9500, Ap
c 18	90	9.3	2910	4	US-09-252-991A-9547	Sequence 9547, Ap
c 19	87.5	9.1	594	4	US-09-252-991A-10010	Sequence 10010, A
c 20	87	9.0	1539	4	US-09-252-991A-1015	Sequence 1015, Ap
c 21	87	9.0	1596	4	US-09-252-991A-1015	Sequence 1015, Ap
c 22	86.5	9.0	2967	4	US-09-185-501B-12	Sequence 12, Appl
c 23	85.5	8.9	990	4	US-09-252-991A-9203	Sequence 9203, Ap
c 24	85.5	8.9	1257	4	US-09-252-991A-16062	Sequence 16062, A
c 25	85.5	8.9	1341	4	US-09-252-991A-8747	Sequence 8747, Ap
c 26	85	8.8	20235	1	US-07-642-734C-3	Sequence 3, Appli
c 27	85	8.8	20235	3	US-08-439-009A-3	Sequence 3, Appli
c 28	84.5	8.7	702	4	US-09-252-991A-4231	Sequence 4231, Ap
c 29	84.5	8.7	963	4	US-09-252-991A-4443	Sequence 4443, Ap
c 30	84.5	8.7	1809	4	US-09-252-991A-9727	Sequence 9727, Ap
c 31	84.5	8.7	3084	4	US-09-252-991A-9645	Sequence 9645, Ap
c 32	84	8.7	43280	2	US-08-804-227C-1	Sequence 1, Appli
c 33	83.5	8.6	1047	4	US-09-252-991A-821	Sequence 821, App
c 34	83.5	8.6	1212	4	US-09-252-991A-807	Sequence 807, App
c 35	83.5	8.6	1215	4	US-09-252-991A-847	Sequence 847, App
c 36	83.5	8.6	1287	4	US-09-252-991A-7292	Sequence 7292, Ap
c 37	83	8.6	2886	4	US-09-221-017B-863	Sequence 863, App
c 38	82	8.5	732	4	US-09-252-991A-12697	Sequence 12697, A
c 39	82	8.5	1260	4	US-09-252-991A-13265	Sequence 13265, A
c 40	82	8.5	3364	2	US-08-735-609-9	Sequence 9, Appli
c 41	82	8.5	3364	2	US-08-735-609-9	Sequence 9, Appli
c 42	82	8.5	3364	3	US-09-315-752-9	Sequence 9, Appli
c 43	82	8.5	3364	3	US-09-244-752-9	Sequence 9, Appli
c 44	82	8.5	3364	3	US-09-245-497-9	Sequence 9, Appli
c 45	82	8.5	3364	4	US-09-562-919-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1

US-09-315-689-4
; Sequence 4, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Folkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-315-689-4

Alignment Scores:
Pred. No.: 7,43e-88 Length: 546
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: 4 Gaps: 0

US-09-938-391-4 (1-184) x US-09-315-689-4 (1-546)

QY	1	HistHrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro	20
Db	1	CACAGCCACCGGACTTCCAGCGGTGCTCCACCTGGTGGCTCAACAGCCCGCTGCA	60
QY	21	GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnAlaArgAlaAla	40
Db	61	GGCGGCATCGGGGATCCGCGGGCGGACTTCCAGTGTCTTCAGCAGCGCGCGGTG	120

Alignment Scores: 1.06e-86 3394
Pred. No.: 835.00 Matches: 154
Score: 92.82% Conservative: 14
Percent Similarity: 92.82% Mismatches: 13
Best Local Similarity: 85.08% Indels: 0
Query Match: 86.44% Gaps: 0
DB: 1

US-09-938-391-4 (1-184) x US-08-159-784-4 (1-3394)

QY 1 HistHrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
DB 1504 CACAGCCACCGGACTTCAGCGGTGCTCCACCTGGTGGCTCAACAGCCCCCTGTCA 1563

QY 21 GlyGlyMetArgGlyVileArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAla 40
DB 1564 GCGGCATCGGGGATCCGCGGGCGGACTTCAGTGTTCACAGCGCGGGCCGTG 1623

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerile 60
DB 1624 GGGCTGGCGGGCACCTTCGCGCCTTCCTGCTCGCGCTGCAGGACCTGTACAGCATC 1683

QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 1684 GTGCGCGGTGCG 1743

QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnGlnLeuLysProGlyAlaArgile 100
DB 1744 CCACGCTGGAGGCTCTGTCTCAGGCTCTGAGGTCCGCTGAGCCCGGGCGCAGCATC 1803

QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
DB 1804 TTCTCTTTGACGGCAAGGACGCTCTGAGGACCCACCTGCGCCCGAGAGGCGTGTG 1863

QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
DB 1864 CATGCTCGGACCCCAACGGGGCGAGGCTGACCGAGAGCTACTGTGAGACGTGGCGGAG 1923

QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuAlaGlyArgLeuGluGln 160
DB 1924 GAGGCTCCCTCGGCACCGGCGCAGGCTCTCGTCTGGGGGCGAGGCTCTGGGGGAG 1983

QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
DB 1984 AGTGGCGCGAGCTGCCATCACCGCTACATCGTCTCTGCTGATTGAGACAGCTTCATGACT 2043

QY 181 Ser 181
DB 2044 GCC 2046

RESULT 4
US-09-449-293-3
; Sequence 3, Application US/09449293
; Patent No. 6267954
; GENERAL INFORMATION:
; APPLICANT: Abitol, Marc
; APPLICANT: Uteza, Yves
; APPLICANT: Menasche, Maurice
; APPLICANT: Bossard, Carine
; APPLICANT: Van Den Bergh, Loic
; APPLICANT: Bonnel, Sebastien
; APPLICANT: Prats, Herve
; APPLICANT: Honiger, Jiri
; APPLICANT: Neuner-Jehle, Martin
; TITLE OF INVENTION: INTRAOCULAR TRANSPLANTATION OF ENCAPSULATED CELLS
; FILE REFERENCE: 8076.202US01
; CURRENT APPLICATION NUMBER: US/09/449,293
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 558
; TYPE: DNA

ORGANISM: Rattus rattus
US-09-449-293-3

Alignment Scores: 2.23e-87 558
Pred. No.: 831.00 Matches: 154
Score: 92.93% Conservative: 17
Percent Similarity: 92.70% Mismatches: 13
Best Local Similarity: 86.02% Indels: 0
Query Match: 86.02% Gaps: 0
DB: 3

US-09-938-391-4 (1-184) x US-09-449-293-3 (1-558)

QY 1 HistHrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
DB 4 CATACTCATCAGGACTTTTCAGCCAGTGTCCACCTGGTGGCAGTCAACACCCCCCTGTCT 63

QY 21 GlyGlyMetArgGlyVileArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAla 40
DB 64 GAGGCGATCGGTGTATCCGTGGAGCAGATTTCAGTGTCTCCAGCAAGCCGAGCCGTG 123

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerile 60
DB 124 GGGCTGTGGGCGACCTTCGGGGCTTCTGCTCTAGGCTGCGAGATCTCTATAGCATC 183

QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 184 GTGCGCGGTGCTGACCGGGGTCTGTGCCATCGTCAACCTGAGGAGCGAGGTGTCTATCT 243

QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnGlnLeuLysProGlyAlaArgile 100
DB 244 CCCAGCTGGGACTCTCTCTTTCTGGCTCCAGGGTCAACTGCAACCCCGGGCGCGCATC 303

QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
DB 304 TTTCTTTTGACGGCAGAGATGCTCTGAGACACCCAGCCTGGCCGAGAGGCGTATGG 363

QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
DB 364 CACGCTCGGACCCAGTGGGGGAGGCTGTGAGAGATGTACTGTGAGACATGGCGAATC 423

QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
DB 424 GAAACTACTGGGGCTACAGGTCAAGCTCAGGCTCTCCCTGCTGTCTGAGGAGGCTCTGGAACAG 483

QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
DB 484 AAGCTGGAGCTGCCCAACAGCTACATCGTCTGTGATTGAGATAGTTCATGACC 543

QY 181 SerPheSerLys 184
DB 544 TCTTCTCCAA 555

RESULT 5
US-09-775-325-3
; Sequence 3, Application US/09775325
; Patent No. 6500449
; GENERAL INFORMATION:
; APPLICANT: Abitol, Marc
; APPLICANT: Uteza, Yves
; APPLICANT: Menasche, Maurice
; APPLICANT: Bossard, Carine
; APPLICANT: Van Den Bergh, Loic
; APPLICANT: Bonnel, Sebastien
; APPLICANT: Prats, Herve
; APPLICANT: Honiger, Jiri
; APPLICANT: Neuner-Jehle, Martin
; TITLE OF INVENTION: INTRAOCULAR TRANSPLANTATION OF ENCAPSULATED CELLS
; FILE REFERENCE: 8076.202US01
; CURRENT APPLICATION NUMBER: US/09/775,325
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: US 09/449,293
; PRIOR FILING DATE: 1999-11-24

; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Rattus rattus
US-09-775-325-3

Alignment Scores:
Pred. No.: 2,23e-87 Length: 558
Score: 831.00 Matches: 154
Percent Similarity: 92.93% Conservative: 17
Best Local Similarity: 83.70% Mismatches: 13
Query Match: 86.02% Indels: 0
DB: 4 Gaps: 0

US-09-938-391-4 (1-184) x US-09-775-325-3 (1-558)

```
QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
Db 4 CATACTCATCAGGACTTTCAGCCAGTGCTCCACCTGGTGCACTGAACACCCCTGTCT 63
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
Db 64 GGAGGCATCGTGTATCCGTGGAGCAGATTTCCAGTGTCTCCAGCAAGCCGCGGTG 123
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerile 60
Db 124 GGGCTGTGGGACCTTCCGGGCTTTCTGTCTCTAGGCTGAGGATCTCTATAGCATC 183
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
Db 184 GTGGCGCGGTGACCGGGGCTGTGGCCCATCGTCAACTGAAGCAGGAGGTGTCTATCT 243
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyAlaArgile 100
Db 244 CCAGCTGGGACTCCCTGTGTTCGGCTCCAGGGTCAACTGCAACCCGGGGCCGCATC 303
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 304 TTTTCTTTTGACGGCAGAGATGCTTGACACCCAGCTGGCCGCGAAGAGCGGTATGG 363
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
Db 364 CACGGCTCGGACCCAGTGGCGGAGGCTGATGAGAGATTACTGTGAGACATGGCGAAT 423
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
Db 424 GAAACTACTGGGGCTACAGTCAGGCTCTCTCCCTGCTGTCAGGCAAGGCTCTGGAACAG 483
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
Db 484 AAAGCTGCGAGCTGCCACACACTACATCATCTCTCTGTGATTGAGAATAGTTCATGACC 543
QY 181 SerPheSerLys 3184
Db 544 TCTTCTCCAAA 555
```

RESULT 6

US-08-985-526-37
; Sequence 37, Application US/08985526
; Patent No. 6080728
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER DNA COMPLEXES CONTAINING DNA
; ; Sequence 37, Application US/08985526
; ; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; ; TITLE OF INVENTION: THERAPY
; ; NUMBER OF SEQUENCES: 43
; ; CORRESPONDENCE ADDRESS:
; ; ADDRESSEE: Connolly, Bove, Lodge, & Hutcz
; ; STREET: 1220 Market Street, P.O. Box 2207
; ; CITY: Wilmington
; ; STATE: Delaware

; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,526
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorrow Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 565 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-985-526-37

Alignment Scores:
Pred. No.: 2,27e-87 Length: 565
Score: 831.00 Matches: 154
Percent Similarity: 92.93% Conservative: 17
Best Local Similarity: 83.70% Mismatches: 13
Query Match: 86.02% Indels: 0
DB: 3 Gaps: 0

US-09-938-391-4 (1-184) x US-08-985-526-37 (1-565)

```
QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
Db 11 CATACTCATCAGGACTTTCAGCCAGTGCTCCACCTGGTGCACTGAACACCCCTGTCT 70
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
Db 71 GGAGGCATCGTGTATCCGTGGAGCAGATTTCCAGTGTCTCCAGCAAGCCGCGGTG 130
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerile 60
Db 131 GGGCTGTGGGACCTTCCGGGCTTTCTGTCTCTAGGCTGAGGATCTCTATAGCATC 190
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
Db 191 GTGCCCGCTGTGACCGGGGCTGTGCCCCTCGTCACCTGAAGCAGGAGGTGTATCT 250
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLeuLysProGlyAlaArgile 100
Db 251 CCAGCTGGGACTCCCTGTGTTCGGCTCCAGGGTCAACTGCAACCCGGGGCCGCATC 310
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 311 TTTTCTTTTGACGGCAGAGATGCTCTGAGACACCCAGGCTGGCCGCGAAGAGCGGTATGG 370
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
Db 371 CACGGCTCGGACCCAGTGGCGGAGGCTGTGAGAGATTACTGTGAGACATGGCGAAT 430
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
Db 431 GAAACTACTGGGGCTACAGGTACAGGTCCCTCCCTGTGTCTGAGGAGGCTCTCTGGAACAG 490
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
Db 491 AAAGCTGCGAGCTGCCACACAGCTACATCATCTCTGTGATTGAGAATAGTTCATGACC 550
```

QY 181 SerPheSerLys 184
Db 551 TCTTTCTCCAAA 562

RESULT 7

US-09-561-500-12

; Sequence 12, Application US/09561500

; Patent No. 6342219

; GENERAL INFORMATION:

; APPLICANT: Philip E. Thorpe

; APPLICANT: Rolf A. Brekken

; TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF

; FILE REFERENCE: 4001.002500

; CURRENT APPLICATION NUMBER: US/09/561,500

; CURRENT FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: 60/131,432

; PRIOR FILING DATE: 1999-04-28

; NUMBER OF SEQ ID NOS: 44

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 12

; TYPE: DNA

; LENGTH: 573

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC

; OTHER INFORMATION: OLIGONUCLEOTIDE

; NAME/KEY: CDS

; LOCATION: (1)..(573)

US-09-561-500-12

Alignment Scores:

Pred. No.:	2,32e-87	Length:	573
Score:	831.00	Matches:	154
Percent Similarity:	92.93%	Conservative:	17
Best Local Similarity:	83.70%	Mismatches:	13
Query Match:	86.02%	Indels:	0
DB:	4	Gaps:	0

US-09-938-391-4 (1-184) x US-09-561-500-12 (1-573)

QY 1 HistHrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
Db 22 CATACTCATCAGGACTTTCCAGCAGTGTCCACCTGGTGCACTGAACACCCCTGTCT 81
QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaA 40
Db 82 GGAGGCATCGGTGATCCGTGGAGCAGATTCCAGTGCTTCAGCAAGCCCGAGCCGTG 141
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerile 60
Db 142 GGGCTGTGGGACCTTCGGGGCTTTCTGTCTCTAGGCTGCAGGATCTCTATAGCATC 201
QY 61 ValArgArgAlaAspArgThrGlyValProValAsnLeuArgAspGluValLeuPhe 80
Db 202 GTGCGCGGTGTGACCGGGGTCTGTGCGCATCGTCAACCTGAAGCAGGAGTGCTATCT 261
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgille 100
Db 262 CCAGCTGGAGATCCCTGTTTCTGGCTCCAGGGTCAACTGCAACCCGGGGCCGCATC 321
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 322 TTTTCTTTGACGGCAGAGATGCTCTGAGACACCCAGACCTGGCGCAGAGCGATGCG 381
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
Db 382 CACGGCTCGACCCCAAGTGGGGAGGCTGATGAGAGGTACTGTGAGACATGGCGAACT 441
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuAlaGlyArgLeuGluGln 160
Db 442 GAAACTACTGGGGCTACAGTACGGCTCTCCCTGTGTGAGGAGGCTCTCTGGAAACAG 501
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180

Db 502 AAAGCTGGAGCTGCCAACACAGCTACATCGTCTGTGCTGAGATAGTCTCATGACC 561
QY 181 SerPheSerLys 184
Db 562 TCTTTCTCCAAA 573

RESULT 8

US-09-561-108-12

; Sequence 12, Application US/09561108

; Patent No. 6342221

; GENERAL INFORMATION:

; APPLICANT: Philip E. Thorpe

; APPLICANT: Rolf A. Brekken

; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF

; FILE REFERENCE: 4001.002584

; CURRENT APPLICATION NUMBER: US/09/561,108

; CURRENT FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: 60/131,432

; PRIOR FILING DATE: 1999-04-28

; NUMBER OF SEQ ID NOS: 44

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 12

; TYPE: DNA

; LENGTH: 573

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC

; OTHER INFORMATION: OLIGONUCLEOTIDE

; NAME/KEY: CDS

; LOCATION: (1)..(573)

US-09-561-108-12

Alignment Scores:

Pred. No.:	2,32e-87	Length:	573
Score:	831.00	Matches:	154
Percent Similarity:	92.93%	Conservative:	17
Best Local Similarity:	83.70%	Mismatches:	13
Query Match:	86.02%	Indels:	0
DB:	4	Gaps:	0

US-09-938-391-4 (1-184) x US-09-561-108-12 (1-573)

QY 1 HistHrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
Db 22 CATACTCATCAGGACTTTCCAGCAGTGTCCACCTGGTGCACTGAACACCCCTGTCT 81
QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaA 40
Db 82 GGAGGCATCGGTGATCCGTGGAGCAGATTCCAGTGCTTCAGCAAGCCCGAGCCGTG 141
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerile 60
Db 142 GGGCTGTGGGACCTTCGGGGCTTTCTGTCTCTAGGCTGCAGGATCTCTATAGCATC 201
QY 61 ValArgArgAlaAspArgThrGlyValProValAsnLeuArgAspGluValLeuPhe 80
Db 202 GTGCGCGGTGTGACCGGGGTCTGTGCGCATCGTCAACCTGAAGCAGGAGTGCTATCT 261
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgille 100
Db 262 CCAGCTGGAGATCCCTGTTTCTGGCTCCAGGGTCAACTGCAACCCGGGGCCGCATC 321
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 322 TTTTCTTTGACGGCAGAGATGCTCTGAGACACCCAGACCTGGCGCAGAGCGATGCG 381
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
Db 382 CACGGCTCGACCCCAAGTGGGGAGGCTGATGAGAGGTACTGTGAGACATGGCGAACT 441
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuAlaGlyArgLeuGluGln 160

Db 442 GAAACTACTGGGCTACAGCTCAGGCTCTCCCTGCTGTCAGGAGGCTCTCTGGAACAG 501
Qy 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
Db 502 AAAGCTGCGAGCTGCCAACACAGCTACATGCTCTGTGCATTGAGAATAGCTTCATGACC 561
Qy 181 SerPheSerLys 184
Db 562 TCTTTCTCCAAA 573
RESULT 9
US-09-561-526-12
; Sequence 12, Application US/09561526
; Patent No. 6416758
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE KITS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002586
; CURRENT APPLICATION NUMBER: US/09/561,526
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; NAME/KEY: CDS
; LOCATION: (1)..(573)
US-09-561-526-12
Alignment Scores:
Pred. No.: 2,32e-87 Length: 573
Score: 831.00 Matches: 154
Percent Similarity: 92.93% Conservative: 17
Best Local Similarity: 83.70% Mismatches: 13
Query Match: 86.02% Indels: 0
Gaps: 0
US-09-938-391-4 (1-184) x US-09-561-526-12 (1-573)
Qy 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
Db 22 CATACTCATCAGGACTTTTCAGCCAGTGTCTCCACCTGGTGGCACTGAACACCCCTGTCT 81
Qy 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
Db 82 GGAGGCGATCGTGGTATCCGTCGAGCAGATTTCCAGTGTCTCCAGCAAGCCGAGCCGTG 141
Qy 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 142 GGGCTGTGGGCACTTCCGGGCTTTCCTGTCTCTAGGCTGAGGATCTCTATAGCATC 201
Qy 61 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
Db 202 GTGCGCGGTGTCACCGGGGCTGTGTGCCATCGTCAACCTGAAGACGAGGTGTATCT 261
Qy 81 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyAlaArgIle 100
Db 262 CCAGCTGGGACTCCCTGTTTCTGGCTCCAGGGTCAACTGCAACCCGGGGCCGCATC 321
Qy 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 322 TTTTCTTTTGACGGCAGAGATGTCCTGAGACACCCAGCTGGCCGCGAGAGCGTATGG 381
Qy 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
Db 382 CACGGCTCGACCCAGTGGCGGAGGCTGATGGAGAGTACTGTGTGAGACATGGCGAACT 441

Qy 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGluGln 160
Db 442 GAAACTACTGGGCTACAGCTCAGGCTCTCCCTGCTGTCAGGAGGCTCTCTGGAACAG 501
Qy 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
Db 502 AAAGCTGCGAGCTGCCAACACAGCTACATGCTCTGTGCATTGAGAATAGCTTCATGACC 561
Qy 181 SerPheSerLys 184
Db 562 TCTTTCTCCAAA 573
RESULT 10
US-09-561-499-12
; Sequence 12, Application US/09561499
; Patent No. 6524583
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002582
; CURRENT APPLICATION NUMBER: US/09/561,499
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; NAME/KEY: CDS
; LOCATION: (1)..(573)
US-09-561-499-12
Alignment Scores:
Pred. No.: 2,32e-87 Length: 573
Score: 831.00 Matches: 154
Percent Similarity: 92.93% Conservative: 17
Best Local Similarity: 83.70% Mismatches: 13
Query Match: 86.02% Indels: 0
Gaps: 0
US-09-938-391-4 (1-184) x US-09-561-499-12 (1-573)
Qy 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
Db 22 CATACTCATCAGGACTTTTCAGCCAGTGTCTCCACCTGGTGGCACTGAACACCCCTGTCT 81
Qy 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
Db 82 GGAGGCGATCGTGGTATCCGTCGAGCAGATTTCCAGTGTCTCCAGCAAGCCGAGCCGTG 141
Qy 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 142 GGGCTGTGGGCACTTCCGGGCTTTCCTGTCTCTAGGCTGAGGATCTCTATAGCATC 201
Qy 61 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
Db 202 GTGCGCGGTGTCACCGGGGCTGTGTGCCATCGTCAACCTGAAGACGAGGTGTATCT 261
Qy 81 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyAlaArgIle 100
Db 262 CCAGCTGGGACTCCCTGTTTCTGGCTCCAGGGTCAACTGCAACCCGGGGCCGCATC 321
Qy 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 322 TTTTCTTTTGACGGCAGAGATGTCCTGAGACACCCAGCTGGCCGCGAGAGCGTATGG 381

QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyxCysGluThrTrpArgThr 140
Db 382 CACGGCTCGACCCAGTGGCGGAGCTGTAGTGGAGTTACTGTGAGACATGGCGAACT 441
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
Db 442 GAAACTACTGGGGCTACAGGTGAGGCTCTCCCTGCTGTGAGGAGGCTCTTGGAAACAG 501
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
Db 502 AAAGTCTCGAGTGCACACACAGCTACATCGTCTGTGATTGAGATAGCTTCATGACC 561
QY 181 SerPheSerLys 184
Db 562 TCTTTTCTCCAAA 573

RESULT 11

US-08-159-784-1
; Sequence 1, Application US/08159784
; Patent No. 5643783
; GENERAL INFORMATION:
; APPLICANT: Bjorn R. Olsen
; TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,784
; FILING DATE: December 1, 1993
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER:
; APPLICATION DATA:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: John F. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00246/170001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4031
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-159-784-1

Alignment Scores:
Pred. No.: 8,79e-86 Length: 4031
Score: 828.00 Matches: 153
Percent Similarity: 92.93% Conservative: 18
Best Local Similarity: 83.15% Mismatches: 13
Query Match: 85.71% Indels: 0
DB: 1 Gaps: 0

US-09-938-391-4 (1-184) x US-08-159-784-1 (1-4031)

QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
Db 3313 CATACTCATCAGGACTTTTCAGGCAGTGTCTCCACCTGTGGTGGCACTGAACACCCCTGTCT 3372

QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
Db 3373 GGAGGCATGCGTGGTATCCGTGGAGCAGATTTCCAGTGTCTCCAGCAAGCCCGAGCCGTG 3432
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrsIle 60
Db 3433 GGGCTGTCTGGGCACCTTCGGGCTTTCTCTAGCTCTAGGCTCAGGATCTCTATAGCATC 3492
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
Db 3493 GTGGCCCGTGTGACCGGGGGTCTGTGCCATCTGTCACCTGAAGACGAGGTGTCTATCT 3552
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLeuPheProGlyAlaArgIle 100
Db 3553 CCCAGCTGGGACTCCCTGTCTTCTGGCTCCAGGGGTCAAGTGCACACCCGGGGCCCGCATC 3612
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 3613 TTTTCTTTTGACGGCAGAGATGTCTGTGACACACCCAGCTGGCCCGCAGAGAGCGTATGG 3672
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyxCysGluThrTrpArgThr 140
Db 3673 CACGGCTCGACCCAGTGGCGGAGGCTGTATGGAGATTACTGTGAGACATGGCGAACT 3732
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
Db 3733 GAAACTACTGGGGCTACAGGTACAGGCTCCCTCCCTGTGTGACGAGGCTCTCTGGAACAG 3792
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
Db 3793 AAAGCTGGGAGTGCACACACAGCTACATCGTCTGTGATTGAGATAGCTTCATGACC 3852
QY 181 SerPheSerLys 184
Db 3853 TCTTTTCTCCAAA 3864

RESULT 12

US-09-315-689-6
; Sequence 6, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Folkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 6
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-315-689-6

Alignment Scores:
Pred. No.: 8,77e-86 Length: 534
Score: 817.00 Matches: 152
Percent Similarity: 92.66% Conservative: 12
Best Local Similarity: 85.88% Mismatches: 13
Query Match: 84.58% Indels: 0
DB: 4 Gaps: 0

US-09-938-391-4 (1-184) x US-09-315-689-6 (1-534)

QY 5 AspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnProGlyGlyMetArg 24
Db 1 GACTTCCAGCCGGTCTCCACCTGTGTGGCTCAACAGCCCTCTGTCAGGCGGCGATGGCG 60

QY 25 GlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaAlaGlyLeuAlaGly 44
Db 61 GGCAATCCGCGGGCGGACTTCCAGTGTCTTCCAGCAGGCGCGGCGCTGGGGCTGGCGGC 120

QY 45 ThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIleValArgAla 64
Db 121 ACCTTCGGCGCCTTCCTGCTCGCGCTGCAGAGACTGTACAGCATGTGCGCGTGC 180
QY 65 AspArgThrGlyValProValValAsnLeuArgAspGluValLeuPheProSerTrpGlu 84
Db 181 GACCGCGCAGCGCTGCCTCATCGTCAAGAGAGAGTGTGTGTTCCTCCAGCTGGAG 240
QY 85 AlaLeuPheSerGlySerGlyGlnLeuValProGlyAlaArgIlePheSerPheAsp 104
Db 241 GCTCTGTCTCAGGCTCTGAGGCTCGCTGAAGCCCGGCGACGATCTTCCTTTGAC 300
QY 105 GlyArgAspValLeuGlnHisProAlaTrpProArgLySerValTrpHisGlySerAsp 124
Db 301 GGCAAGGAGCTCTGAGGACCCACCTGCGCCCGAGAGCGTGTGGCATGGCTCGGAC 360
QY 125 ProSerGlyArgArgIleThrAspSerTyrCysGluThrTrpArgThrGluAlaProAla 144
Db 361 CCCAAGCGCGCAGGCTGACCGAGAGCTACTGTGAGACGCTGGCGGACGGAGGCTCCCTCG 420
QY 145 AlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGlnGluAlaSer 164
Db 421 GCCACGGCGCAGGCTCTCGCTGCTGGGGGCGAGGCTCTGGGGGCGAGTGCCTCGGAG 480
QY 165 CysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThrSer 181
Db 481 TGCATCAGCGCTACATCGTCTGCTGATGAGACAGCTTCATGACTGCC 531

RESULT 13

US-09-252-991A-15455
; Sequence 15455, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15455
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15455

Alignment Scores:
Pred. No.: 0.156 Length: 591
Score: 90.50 Matches: 55
Percent Similarity: 38.71% Conservative: 17
Best Local Similarity: 29.57% Mismatches: 62
Query Match: 9.37% Indels: 53
DB: 4 Gaps: 9

US-09-938-391-4 (1-184) x US-09-252-991A-15455 (1-591)

QY 18 ProGlnProGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAla 37
Db 20 CCTCGCGGGCGGAGCTGCAGCAGCGCGGCGAGGATCGCGG-----AGTGGCGG 67
QY 38 ArgAlaAla-----GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 54
Db 68 CGCGCAGGTCTATGGGAGAGAGCGCAGGAGGCTGGCGCTGGCGCTCGGCATCGC--- 124
QY 55 GlnAspLeuTyrSerIleValArgAlaAspArgThrGlyValPro-----70
Db 125 -----GTTCCGCGCTCAGTGCATCAGGCTCACCAGCCAGCGCATGG 166
QY 71 -----ValValAsn 73

Db 167 CGGTTTCGCTTTTCGACCGCGCAGCGCGCCAGGTTCGAGGCGCAGGTTCGCGTTCGCCGT 226
QY 74 LeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSer-----Gly 89
Db 227 GGTCTCGGGAACTCGATGGCTTCACAGGTTTCAGCGTTTGGCGGGCGCCACCGCGGAGGC 286
QY 90 SerGluGlyGlnLeuLyProGlyAla-----ArgIlePheSerPheAspGlyArg 106
Db 287 TCGATGGCGGAGTTCCGCCAGCGCTGGCGCAGTTCCGGCGCGCGCAGCTCGCGCCGA 346
QY 107 AspValLeuGlnHisProAlaTrpProArgLySerValTrpHisGlySerAspProSer 126
Db 347 ACAGGTTCGCGTCCCTTCGCTGGTTGGCCAGCA-----GGATCTGCCCTCGC 391
QY 127 GlyArgArgLeuThrAspSer-TyrCysGluThrTrpArgThr-----GluAlaPr 143
Db 392 GGTCTCAGCA---TCAGGCTCGCCAGCGCAGCATTCACAGCGCTCGGCCATGAACCGCC 448
QY 143 AlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGluGlnGluAla 163
Db 449 GGCTGTCCGGGTGGCGCTGACCGCTGCTCCA-GGGCGCGCATGGGTTCTGCAATGGA 507
QY 163 aSerCysArgHisAla 168
Db 508 TCGTTGCCGCGCGCGC 523

RESULT 14

US-09-252-991A-15284/c
; Sequence 15284, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15284
; LENGTH: 2319
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15284

Alignment Scores:
Pred. No.: 1.14 Length: 2319
Score: 90.50 Matches: 55
Percent Similarity: 38.71% Conservative: 17
Best Local Similarity: 29.57% Mismatches: 62
Query Match: 9.37% Indels: 53
DB: 4 Gaps: 9

US-09-938-391-4 (1-184) x US-09-252-991A-15284 (1-2319)

QY 18 ProGlnProGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAla 37
Db 1741 CCTCGCGGGCGGAGCTGCAGCAGCGCGGCGAGGATCGCGG-----AGTGGCGG 1694
QY 38 ArgAlaAla-----GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 54
Db 1693 CGCGCAGGTCTATGGGAGAGAGCGCAGGAGGCTGGCGCTGGCGCTCGGCATCGC--- 1637
QY 55 GlnAspLeuTyrSerIleValArgAlaAspArgThrGlyValPro-----70
Db 1636 -----GTTCCGCGCTCAGTGCATCAGGCTCACCAGCCAGCGCATGG 1595
QY 71 -----ValValAsn 73

Db 1594 CGGTTTCGCTTCCACCGCAGCAGCGCGCAGGTCGAGCGCGCAGGTCGCGTTCGCCGT 1535
QY 74 LeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSer-----Gly 89
Db 1534 GGTCCGGAACTCGATGGCTTCAGGGTTCCAGCGTTGGCGGGCGCCACCGCGGAGC 1475
QY 90 SerGluGlyGlnLeuLysProGlyAla-----ArgIlePheSerPheAspGlyArg 106
Db 1474 TCGGATGGCGGAGTTTCGGCGCAGCGCTCGCGCAGTTCCGCGCGCGCAGCTCGCGCGCA 1415
QY 107 AspValLeuGlnHisProAlaTrpProArgLysSerValTrpHisGlySerAspProSer 126
Db 1414 ACAGTTCGCGTTCGCTTCGCGTTCGCGCAGCA-----GGATCTGCCCTTCGCG 1370
QY 127 GlyArgArgLeuThrAspSer-TyrCysGluThrTrpArgThr-----GluAlaPr 143
Db 1369 GGTTCGCGA---TCAGGTTCGCGCAGCGCAGGATTCACAGCGCTCGCGCATGAACCGCC 1313
QY 143 oAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGlnGluAlaAl 163
Db 1312 GGCTGTTCGCGGTTCGCGCTGACCGCTGCTCCA-GGGCGCGCATGCGGTTCTGCAATGGA 1254
QY 163 aSerCysArgHisAla 168
Db 1253 TCGTTGCCGCGCGCGC 1238

RESULT 15

US-09-252-991A-15396
; Sequence 15396, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15396
; LENGTH: 2427
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15396

Alignment Scores:

Pred. No.:	1.21	Length:	2427
Score:	90.50	Matches:	55
Percent Similarity:	38.71%	Conservative:	17
Best Local Similarity:	29.57%	Mismatches:	62
Query Match:	9.37%	Indels:	53
DB:	4	Gaps:	9

US-09-938-391-4 (1-184) x US-09-252-991A-15396 (1-2427)

QY 18 ProGlnProGlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAla 37
Db 393 CCTCGCGCGGCGGACGTCGACGAGCGCGGATCGCGG-----AGTGGCGCG 440
QY 38 ArgAlaAla-----GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeu 54
Db 441 CGCGGAGGTCATGGAGAGAGCGCAGGAGGGTGGCGGCTCGCGCATCGC--- 497
QY 55 GlnAspLeuTyrSerIleValArgArgAlaAspArgThrGlyValPro----- 70
Db 498 -----GTTTCGCGCTCAGGTCAGTTCAGGCTCACCAGCCAGCGATCG 539
QY 71 -----ValValAsn 73
Db 540 CGGTTTCGCTTTCACCGCGCAGCAGCGCGCGCAGGTCGAGCGCGCAGGTCGCGTTCGCCGT 599

QY 74 LeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSer-----Gly 89
Db 600 GGTCCGGAACTCGATGGCTTCAGGGTTTCAGCGTTGGCGGGCGCCACCGCGGAGC 659
QY 90 SerGluGlyGlnLeuLysProGlyAla-----ArgIlePheSerPheAspGlyArg 106
Db 660 TCGGATGGCGGAGTTTCGGCGCAGCGCTCGCGCAGTTCCGCGCGCGCAGCTCGCGCGCA 719
QY 107 AspValLeuGlnHisProAlaTrpProArgLysSerValTrpHisGlySerAspProSer 126
Db 720 ACAGTTCGCGTTCGCTTCGCGTTCGCGCAGCA-----GGATCTGCCCTTCGCG 764
QY 127 GlyArgArgLeuThrAspSer-TyrCysGluThrTrpArgThr-----GluAlaPr 143
Db 765 GGTTCGCGA---TCAGGTTCGCGCAGCGCAGGATTCACAGCGCTCGCGCATGAACCGCC 821
QY 143 oAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGlnGluAlaAl 163
Db 822 GGCTGTTCGCGGTTCGCGCTGACCGCTGCTCCA-GGGCGCGCATGCGGTTCTGCAATGGA 880
QY 163 aSerCysArgHisAla 168
Db 881 TCGTTGCCGCGCGCGC 896

Search completed: August 17, 2003, 23:00:11
Job time : 62.7778 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 17, 2003, 21:56:45 ; Search time 215.556 Seconds
(without alignments)
1910.324 Million cell updates/sec

Title: us-09-938-391-4

Perfect score: 966

Sequence: 1 HTHQDFQLVHLVALNSQP.....CRHAFVVLCIENSVMTSFK 184

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1504479 seqs, 1118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n model -DEV=xlp

-Q/cgn2_1/USFTO_spool_p/US0938391/runat_04082003_130658_27206/app_query.fasta_1.718
-DB=Published Applications NA -QWTF=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USFR=US0938391 @CGN 1.1 129 @runat_04082003_130658_27206
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	959	99.3	552 12	US-10-292-418-34

2	959	99.3	552 14	US-10-131-241-50	Sequence 50, Appl
3	891	92.2	632 14	US-10-131-241-51	Sequence 51, Appl
4	835	86.4	546 14	US-10-042-347-4	Sequence 4, Appl
5	835	86.4	549 12	US-10-292-418-3	Sequence 3, Appl
6	835	86.4	549 14	US-10-131-241-53	Sequence 53, Appl
7	835	86.4	552 9	US-09-873-676-30	Sequence 30, Appl
8	835	86.4	3394 10	US-09-880-107-2178	Sequence 2178, Ap
9	835	86.4	4551 14	US-10-060-036-144	Sequence 144, App
10	831	86.0	540 14	US-10-131-241-48	Sequence 48, Appl
11	831	86.0	558 9	US-09-775-174-3	Sequence 3, Appl
12	831	86.0	558 9	US-09-775-325-3	Sequence 3, Appl
13	831	86.0	565 13	US-10-036-869-37	Sequence 37, Appl
14	831	86.0	573 10	US-09-998-831-12	Sequence 12, Appl
15	828	85.7	552 12	US-10-232-418-17	Sequence 17, Appl
16	828	85.7	624 13	US-10-080-797-4	Sequence 4, Appl
17	827	85.6	551 13	US-10-080-797-2	Sequence 2, Appl
18	817	84.6	534 14	US-10-042-347-6	Sequence 6, Appl
19	817	84.6	537 14	US-10-131-241-59	Sequence 59, Appl
20	634	65.6	900 14	US-10-131-241-45	Sequence 45, Appl
21	526	54.5	5204 12	US-09-971-392-159	Sequence 159, App
22	358	37.1	574 14	US-10-060-036-64	Sequence 64, Appl
23	151	15.6	434 14	US-10-060-036-4	Sequence 4, Appl
24	100.5	10.4	1344 9	US-09-815-242-7650	Sequence 7650, Ap
25	98	10.1	9025608 14	US-10-156-761-1	Sequence 1, Appl
26	93.5	9.7	1656 13	US-10-027-632-37643	Sequence 97643, A
27	90.5	9.4	713 13	US-10-027-632-142192	Sequence 142192,
28	90	9.3	251 10	US-09-880-107-713	Sequence 713, App
29	89	9.2	2439 14	US-10-156-761-2569	Sequence 2569, Ap
30	86	8.9	1465 9	US-09-864-761-9885	Sequence 9885, Ap
31	85.5	8.9	22118 10	US-09-799-4624-16	Sequence 16, Appl
32	85.5	8.9	22118 11	US-09-815-981-5	Sequence 5, Appl
33	85.5	8.9	22118 11	US-09-836-911A-16	Sequence 16, Appl
34	85.5	8.9	22118 11	US-09-815-979-5	Sequence 5, Appl
35	85.5	8.9	22118 13	US-10-125-767-16	Sequence 16, Appl
36	85.5	8.9	22118 14	US-10-151-081-16	Sequence 16, Appl
37	85.5	8.9	22118 14	US-10-287-313-16	Sequence 16, Appl
38	85.5	8.9	22118 14	US-10-219-694-16	Sequence 16, Appl
39	85.5	8.9	22118 14	US-10-235-119-5	Sequence 5, Appl
40	85.5	8.9	22118 14	US-10-161-403-18	Sequence 18, Appl
41	83.5	8.6	1703 14	US-10-225-567A-309	Sequence 309, App
42	83.5	8.6	15952 14	US-10-171-311-51	Sequence 51, Appl
43	83.5	8.6	173808 13	US-10-003-806-10	Sequence 10, Appl
44	83	8.6	181 10	US-09-867-701-3816	Sequence 3816, Ap
45	83	8.6	13483 12	US-10-017-161-2071	Sequence 2071, Ap

ALIGNMENTS

RESULT 1
US-10-292-418-34
; Sequence 34, Application US/10292418
; Publication No. US20030139365A1
; GENERAL INFORMATION:
; APPLICANT: Lo, Kin-Ming
; APPLICANT: Li, Yue
; APPLICANT: Gillies, Stephen D
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
; TITLE OF INVENTION: Immunofusins
; FILE REFERENCE: LEX-006C1
; CURRENT APPLICATION NUMBER: US/10/292,418
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/383,315
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: US 60/097,883
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS

; LOCATION: (1)..(552)
; OTHER INFORMATION: Endostatin
US-10-292-418-34

Alignment Scores: 1,45e-108 Length: 552
Pred. No.: 959.00 Matches: 183
Score: 99.46% Conservative: 0
Percent Similarity: 99.46% Mismatches: 1
Best Local Similarity: 99.46% Indels: 0
Query Match: 99.28% Gaps: 0
DB: 12%

US-09-938-391-4 (1-184) x US-10-292-418-34 (1-552)

QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
DB 1 CACACCCACAGGACTTCCAGCGCGTGTGCACCTGGTGGCCCTGAACAGCGCGAGCGG 60
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnAlaArgAlaAla 40
DB 61 GCGCGCATCGGAGGCAATCCGGGAGCGGACTTCCAGTGTTCAGGAGCGCGCGCGG 120
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerile 60
DB 121 GGGCTGGCGGACCTTCCGGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 181 GTGCGCGCGCGGACCGGCGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgille 100
DB 241 CCCAGCTGGAGGCTTATCTCGGGCTCGAGGCGGCGAGTGAAGCCCGGGCGCGCATC 300
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
DB 301 TTCTCTTTTCGACGAGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
DB 361 CACGGCTCCGACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
DB 421 GAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
DB 481 GAGGCGCGGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
QY 181 SerPheSerLys 184
DB 541 TCCTTCTCCAAG 552

RESULT 2

US-10-131-241-50
; Sequence 50, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT FILING DATE: 2002-07-22
; PRIOR FILING DATE: 1999-10-06
; PRIOR FILING DATE: 1999-05-21
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Canine sp.
US-10-131-241-50

Alignment Scores: 1,45e-108 Length: 552
Pred. No.: 959.00 Matches: 183
Score: 99.46% Conservative: 0
Best Local Similarity: 99.46% Mismatches: 1
Query Match: 99.28% Indels: 0
DB: 14

US-09-938-391-4 (1-184) x US-10-131-241-50 (1-552)

QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
DB 1 CACACCCACAGGACTTCCAGCGCGTGTGCACCTGGTGGCCCTGAACAGCGCGAGCGG 60
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnAlaArgAlaAla 40
DB 61 GCGCGCATCGGAGGCAATCCGGGAGCGGACTTCCAGTGTTCAGGAGCGCGCGCGG 120
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerile 60
DB 121 GGGCTGGCGGACCTTCCGGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 181 GTGCGCGCGCGGACCGGCGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgille 100
DB 241 CCCAGCTGGAGGCTTATCTCGGGCTCGAGGCGGCGAGTGAAGCCCGGGCGCGCATC 300
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
DB 301 TTCTCTTTTCGACGAGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
DB 361 CACGGCTCCGACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
DB 421 GAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
DB 481 GAGGCGCGGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
QY 181 SerPheSerLys 184
DB 541 TCCTTCTCCAAG 552

RESULT 3

US-10-131-241-51
; Sequence 51, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT FILING DATE: 2002-07-22
; PRIOR FILING DATE: 1999-10-06
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65

Db 541 GCC 543

RESULT 5

US-10-292-418-3

; Sequence 3, Application US/10292418

; Publication No. US20030139365A1

; GENERAL INFORMATION:

; APPLICANT: Lo, Kin-Ming

; APPLICANT: Li, Yue

; APPLICANT: Gillies, Stephen D

; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as

; FILE OF INVENTION: Immunofusins

; FILE REFERENCE: LEX-006C1

; CURRENT APPLICATION NUMBER: US/10/292,418

; CURRENT FILING DATE: 2002-11-12

; PRIOR APPLICATION NUMBER: 09/383,315

; PRIOR FILING DATE: 1999-08-25

; PRIOR APPLICATION NUMBER: US 60/097,883

; PRIOR FILING DATE: 1998-08-25

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 549

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(549)

; OTHER INFORMATION: endostatin

US-10-292-418-3

Alignment Scores:

Pred. No.: 2,42e-93 Length: 549

Score: 835.00 Matches: 154

Percent Similarity: 92.82% Conservative: 14

Best Local Similarity: 85.08% Mismatches: 13

Query Match: 86.44% Indels: 0

DB: 12 Gaps: 0

US-09-938-391-4 (1-184) x US-10-292-418-3 (1-549)

QY 1 HisThrHisGlnAspPheGlnLeuValLeuValAlaLeuAsnSerProGlnPro 20

Db 1 CACAGCACCGCGACTTCCAGCGCGTGTCCACCTGCTGCGCTCAACAGCCCCCTGTCA 60

QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40

Db 61 GCGGCGATGCGGGGCGATCCCGGGGCGGACTTCCAGTGTCTCCAGCAGCGCGGGCGGTG 120

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60

Db 121 GGCGTGGCGGCGACTTCCCGCGCTTCTCTCGCGCTGCGAGCTGTACAGCATC 180

QY 61 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80

Db 181 GTGCGCGTCCGACCGCGCAGCGTGTCCCATCGTCAACCTCAAGCAGAGCTGTCTTT 240

QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 100

Db 241 CCCAGCTGGGAGGCTCTGTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGGCGACATC 300

QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTyr 120

Db 301 TTCTCTTTCAGCGCAGGAGGCTCTGAGGACCCACCTGCGCCCGAGAGAGCGTGTGG 360

QY 121 HisGlySerAspProSerGlyArgLeuThrAspSerTyrCysGluThrTrpArgThr 140

Db 361 CATGGCTCGGACCCCAACGCGGCGAGGCTGTACAGAGCTACTGTGAGACGTGGCGGACG 420

QY 141 GluAlaProAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160

Db 421 GAGGCTCCCTCGGCGCAGGCGCTCTCTGCTGCGGGGCGAGGCTCTCTGGGGCAG 480

QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180

Db 481 AGTGGCGGAGCTGCCATCAGCCCTACATCGTCTGCTGATTGAGACAGCTTCATGACT 540

QY 181 Ser 181

Db 541 GCC 543

RESULT 6

US-10-131-241-53

; Sequence 53, Application US/10131241

; Publication No. US20030012792A1

; GENERAL INFORMATION:

; APPLICANT: Holaday, John W.

; APPLICANT: Fortier, Anne H.

; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation

; FILE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers

; FILE REFERENCE: 05213-0344 43170-271565

; CURRENT APPLICATION NUMBER: US/10/131,241

; CURRENT FILING DATE: 2002-07-22

; PRIOR APPLICATION NUMBER: US 09/413,049

; PRIOR FILING DATE: 1999-10-06

; PRIOR APPLICATION NUMBER: US 09/316,802

; PRIOR FILING DATE: 1999-05-21

; PRIOR APPLICATION NUMBER: US 60/086,586

; PRIOR FILING DATE: 1998-05-22

; NUMBER OF SEQ ID NOS: 65

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 53

; LENGTH: 549

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-131-241-53

Alignment Scores:

Pred. No.: 2,42e-93 Length: 549

Score: 835.00 Matches: 154

Percent Similarity: 92.82% Conservative: 14

Best Local Similarity: 85.08% Mismatches: 13

Query Match: 86.44% Indels: 0

DB: 12 Gaps: 0

US-09-938-391-4 (1-184) x US-10-131-241-53 (1-549)

QY 1 HisThrHisGlnAspPheGlnLeuValLeuValAlaLeuAsnSerProGlnPro 20

Db 1 CACAGCACCGCGACTTCCAGCGCGTGTCCACCTGCTGCGCTCAACAGCCCCCTGTCA 60

QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40

Db 61 GCGGCGATGCGGGGCGATCCCGGGGCGGACTTCCAGTGTCTCCAGCAGCGCGGGCGGTG 120

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60

Db 121 GGCGTGGCGGCGACTTCCCGCGCTTCTCTCGCGCTGCGAGCTGTACAGCATC 180

QY 61 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80

Db 181 GTGCGCGTGGCGAGGCTCTGTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGGCGACATC 300

QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 100

Db 241 CCCAGCTGGGAGGCTCTGTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGGCGACATC 300

QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTyr 120

Db 301 TTCTCTTTCAGCGCAGGAGGCTCTGAGGACCCACCTGCGCCCGAGAGAGCGTGTGG 360

QY 121 HisGlySerAspProSerGlyArgLeuThrAspSerTyrCysGluThrTrpArgThr 140

Db 361 CATGGCTCGGACCCCAACGCGGCGAGGCTGTACAGAGCTACTGTGAGACGTGGCGGACG 420

QY 141 GluAlaProAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160

Db 421 GAGGCTCCCTCGGCCACCGGCCAGGCTCTCTGCTGCGGGGCGCAGGCTCTCTGGGGCGAG 480
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
Db 481 AGTGGCGGAGCTGCCATCAGCCTACATCGTCTCTGCATTGAGAACAGCTTCATGACT 540
QY 181 Ser 181
Db 541 GCC 543

RESULT 7

US-09-873-676-30
; Sequence 30, Application US/09873676
; Patent No. US2002007289A1
; GENERAL INFORMATION:
; APPLICANT: Macdonald, Nicholas J.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use
; FILE REFERENCE: 05213-0378 (43170-259333)
; CURRENT APPLICATION NUMBER: US/09/873,676
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 30
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-676-30

Alignment Scores:
Pred. No.: 2,44e-93 Length: 552
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: Gaps: 0

US-09-938-391-4 (1-184) x US-09-873-676-30 (1-552)

QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisValAlaLeuAsnSerProGlnPro 20
Db 1 CACAGCACCGCGACTTCCAGCGGTGCTCCACTGTGGTGGCTCAACAGCCCCCTGTCA 60
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
Db 61 GCGCGCATGCGGGCATTCGCGGGCGGACTTCCAGTGTCTCCAGCAGCGCGGGCGGTG 120
QY 41 GlyValAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTySerIle 60
Db 121 GGCGCTGGCGGACCTTCCGCGCTCTCTGCTCCGCGCTCCAGTGTCTCCAGCAGCTGTACAGCATC 180
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
Db 181 GTGCGCGGTGCCAGCGCGCGCGCGCTCCCTCAACCTCAAGCAGAGCTGTCTGTTT 240
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 100
Db 241 CCCAGCTGGGAGGCTCTGTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGGCGCGCATC 300
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 301 TTCTCTTTGACGGCAGGACGCTCTCAGGACACCCACCTGGCCCCCAGAGAGCGGTGTG 360
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrcysGluThrTrpArgThr 140
Db 361 CATGGCTCGACCCCAACGGCGCGCGCGCTGACCGAGAGCTACTGTGAGACGTGGCGGACG 420
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160

Db 421 GAGGCTCCCTCGGCCACCGGCCAGGCTCTCTGCTGCGGGGCGCAGGCTCTCTGGGGCGAG 480
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
Db 481 AGTGGCGGAGCTGCCATCAGCCTACATCGTCTCTGCATTGAGAACAGCTTCATGACT 540
QY 181 Ser 181
Db 541 GCC 543

RESULT 8

US-09-880-107-2178
; Sequence 2178, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherff, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2178
; LENGTH: 3394
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L22548
US-09-880-107-2178

Alignment Scores:
Pred. No.: 2,39e-92 Length: 3394
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: Gaps: 0

US-09-938-391-4 (1-184) x US-09-880-107-2178 (1-3394)

QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisValAlaLeuAsnSerProGlnPro 20
Db 1504 CACAGCACCGCGACTTCCAGCGGTGCTCCACTGTGGTGGCTCAACAGCCCCCTGTCA 1563
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
Db 1564 GCGCGCATGCGGGCATTCGCGGGCGGACTTCCAGTGTCTCCAGCAGCGCGGGCGGTG 1623
QY 41 GlyValAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTySerIle 60
Db 1624 GGCGCTGGCGGGACCTTCCGCGCTCTCTGCTCCGCGCTGCGAGACCTGTACAGCATC 1683
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
Db 1684 GTGCGCGGTGGCGCGCGCGCGCTGCGCTCAACCTCAAGCAGAGCTGTCTGTTT 1743
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 100
Db 1744 CCGAGCTGGAGGCTCTGTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGGCGCGCATC 1803
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 1804 TTCTCTTTGACGGCAGGACGCTCTCAGGACACCCACCTGGCCCCCAGAGAGCGGTGTG 1863
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrcysGluThrTrpArgThr 140

```

1864 CATGGCTCGGACCCCAACGGCGCGAGGCTGCTGAGAGCTGCGGACG 1923
141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
1924 GAGGCTCCCTCGGCGCACGGCGCGAGGCTCTCTGCTGCTGGGGGCGAGGCTCTGGGGCAG 1983
161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
1984 AGTGGCGGAGCTGCCATCAGCGCTACATCGTCTCTGCATTGAGAACACGCTTCATGACT 2043
181 Ser 181
2044 GCC 2046

RESULT 9
US-10-060-036-144
; Sequence 144, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael J.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 144
; LENGTH: 4551
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-036-144

Alignment Scores:
Pred. No.: 3,46e-92 Length: 4551
Score: 835.00 Matches: 154
Percent Similarity: 92.82% Conservative: 14
Best Local Similarity: 85.08% Mismatches: 13
Query Match: 86.44% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-4 (1-184) x US-10-060-036-144 (1-4551)
QY 1 HisThrHisGlnAspPheGlnLeuValValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
DB 4000 CACAGCACCGGACCTTCAGCGGTGCTCACTGTTGGCTCACAGCCCGCTGTCA 4059
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
DB 4060 GCGGCGATCGGGGCGGATCCCGGGCGGACTTCAGTGCTTCAGCAGGCGGCGCGTG 4119
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
DB 4120 GGCGTGGCGGCGACCTTCGCGGCTTCCTGCTCGCGCTGCGAGGACCTGTACAGCATC 4179
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 4180 GTGGCGCGTCCGACCGCGCGAGCGCTGCTCACTCAAGCAGAGCTGTGTTT 4239
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 100
DB 4240 CCCAGCTGGGAGGCTCTGTTCTCAGGCTCTCAGGCTCCGCTGAAGCCCGGGCGGCGCATC 4299
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
DB 4300 TTCTCTTTGACGCAAGAGCGCTCTGAGGACACCCACCTGGGCCCCCAGAGAGCGGTGG 4359
```

```

121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
4360 CATGGCTCGGACCCCAACGGCGCGAGGCTGACCGAGAGCTACTGTGAGACGTCGCGGACG 4419
141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
4420 GAGGCTCCCTCGGCGCACGGCGCGAGGCTCTCTGCTGCTGGGGGCGAGGCTCTGGGGCAG 4479
161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
4480 AGTGGCGGAGCTGCCATCAGCGCTACATCGTCTCTGCATTGAGAACACGCTTCATGACT 4539
181 Ser 181
4540 GCC 4542

RESULT 10
US-10-131-241-48
; Sequence 48, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 540
; TYPE: DNA
; ORGANISM: Rhesus monkey
US-10-131-241-48

Alignment Scores:
Pred. No.: 7,35e-93 Length: 540
Score: 831.00 Matches: 152
Percent Similarity: 92.22% Conservative: 14
Best Local Similarity: 84.44% Mismatches: 14
Query Match: 86.02% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-4 (1-184) x US-10-131-241-48 (1-540)
QY 1 HisThrHisGlnAspPheGlnLeuValValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
DB 1 CACAGCACCGGACCTTCAGCGGTGCTCACTGTTGGCTCAATAGCCCGCTGCCA 60
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
DB 61 GCGGCGATCGGGGCGGATCCCGGGCGGACTTCAGTGCTTCAGCAGGCGGCGCGTG 120
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
DB 121 GGCGTGGTGGGCGACCTTCGCTGCTCTCAGCGCTGCGAGGACCTGTACAGCATC 180
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 181 GTGGCGCGTGGCGCGAGGCTGCTGAGGCTCCGCTGAGGCTGAGGCTGAGTGTGTTT 240
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 100
DB 241 CCCAGCTGGAGGCTTGTTCGAGGCTCTGAGGCTCCGCTGAGCCCGGGCGGCGCATC 300
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
```

```
Db 301 TTCTCTTTGACGGCAGGACGTCCTGAGGACACCCACCTGGCCCGCAGAGAGCGTGTGG 360
QY 121 HisGlySerAspProSerGlyArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
Db 361 CATGGCTCGGACCCAGCGGCCGACGCTGACTGAGAGTACTGCGAGACGTGGCGGACA 420
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
Db 421 GAGTCTCCCTCGGTCAAGGTCAAGGCTCTCTCCCTGCTGGGGGCGAGGCTCTTAGGGCAG 480
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
Db 481 AATGCCCGCAAGCTGTCAACACGCTATATGCTCTCTGTCATCGAACAAGCTTCATGACT 540

RESULT 11
US-09-775-174-3
; Sequence 3, Application US/09775174
; Patent No. US20020076396A1
; GENERAL INFORMATION:
; APPLICANT: Abitol, Marc
; APPLICANT: Uteza, Yves
; APPLICANT: Menasche, Maurice
; APPLICANT: Bossard, Carine
; APPLICANT: Van Den Bergh, Loic
; APPLICANT: Bonnel, Sebastian
; APPLICANT: Prats, Hervé
; APPLICANT: Honiger, Jiri
; APPLICANT: Neuner-Jehle, Martin
; TITLE OF INVENTION: INTRAOCULAR TRANSPLANTATION OF ENCAPSULATED CELLS
; FILE REFERENCE: 8076.202USD
; CURRENT APPLICATION NUMBER: US/09/775,174
; CURRENT FILING DATE: 2001-02-01
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Rattus rattus
US-09-775-174-3

Alignment Scores:
Pred. No.: 7.66e-93 Length: 558
Score: 831.00 Matches: 154
Percent Similarity: 92.93% Conservative: 17
Best Local Similarity: 83.70% Mismatches: 13
Query Match: 86.02% Indels: 0
DB: 9 Gaps: 0

US-09-938-391-4 (1-184) x US-09-775-174-3 (1-558)
QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
Db 4 CATACTCATCAGGACTTTCAGGCAGTGTCTCCACCTGGTGCACTGAACACCCCTGTCT 63
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAaAla 40
Db 64 GGAGGATCGGTGGTATCCGTGGAGCAGATTTCAGTGTCTTCAGCAAGCCCGCGGTG 123
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 124 GGCTGTGGGCACTTCCGGGCTTCTCTAGGCTGAGGATCTCTATAGCATC 183
QY 61 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
Db 184 GTGGCGGTCTGACCGGGGTCTGTGCCCATCGTCAACCTGAAGACGAGGTGTCTATCT 243
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGluGlnLeuLysProGlyAlaArgIle 100
Db 244 CCAGCTGGGACTCCCTGTTTCTGGCTCCAGGGTCAACTGCAACCCGGGGCCGCATC 303
```

```
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 304 TTTTCTTTTGACGGCAGAGATGCTCTGAGACACCCAGCCTGGCCGAGAGAGCGTATGG 363
QY 121 HisGlySerAspProSerGlyArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
Db 364 CACGGCTCGGACCCAGTGGCGGAGCTGTGAGAGATTACTGTGAGACATGGCGAAT 423
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
Db 424 GAACTACTGGGGCTACAGTCAAGGCTCTCCCTGCTGTCAGGACGGCTCTTGGAACAG 483
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
Db 484 AAAGTGGCGAGCTGCCACAACAGCTACATCGTCTCTGTGCATTGAGAATAGCTTCATGACC 543
QY 181 SerPheSerLys 184
Db 544 TCTTTCTCCAAA 555

RESULT 12
US-09-775-325-3
; Sequence 3, Application US/09775325
; Patent No. US20020076397A1
; GENERAL INFORMATION:
; APPLICANT: Abitol, Marc
; APPLICANT: Uteza, Yves
; APPLICANT: Menasche, Maurice
; APPLICANT: Bossard, Carine
; APPLICANT: Van Den Bergh, Loic
; APPLICANT: Bonnel, Sebastian
; APPLICANT: Prats, Hervé
; APPLICANT: Honiger, Jiri
; APPLICANT: Neuner-Jehle, Martin
; TITLE OF INVENTION: INTRAOCULAR TRANSPLANTATION OF ENCAPSULATED CELLS
; FILE REFERENCE: 8076.202USD
; CURRENT APPLICATION NUMBER: US/09/775,325
; CURRENT FILING DATE: 2001-02-01
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Rattus rattus
US-09-775-325-3

Alignment Scores:
Pred. No.: 7.66e-93 Length: 558
Score: 831.00 Matches: 154
Percent Similarity: 92.93% Conservative: 17
Best Local Similarity: 83.70% Mismatches: 13
Query Match: 86.02% Indels: 0
DB: 9 Gaps: 0

US-09-938-391-4 (1-184) x US-09-775-325-3 (1-558)
QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
Db 4 CATACTCATCAGGACTTTCAGGCAGTGTCTCCACCTGGTGCACTGAACACCCCTGTCT 63
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAaAla 40
Db 64 GGAGGATCGGTGGTATCCGTGGAGCAGATTTCAGTGTCTTCAGCAAGCCCGCGGTG 123
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 124 GGCTGTGGGCACTTCCGGGCTTCTCTAGGCTGAGGATCTCTATAGCATC 183
QY 61 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
Db 184 GTGGCGGTCTGACCGGGGTCTGTGCCCATCGTCAACCTGAAGACGAGGTGTCTATCT 243
```

Qy 81 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyAlaArgIle 100
Db 244 CCAGCTGGGACCTCCCTGTTTCTGGCTCCAGGTCACCTGCAACCCGGGGCCCGCATC 303
Qy 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 304 TTTTCTTTGACGGCAGAGATGCTTGAGACACCCAGCTGGCGCAGAGAGCGTATGG 363
Qy 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
Db 364 CACGGCTCGGACCCAGTGGCGGAGGCTGATGAGAGTACTGTGAGACATGGCGAACT 423
Qy 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
Db 424 GAACTACTGGGCTACAGCTCAGGCTCTCTCTGCTGTCAGCAGGCTCCCTGGAAACAG 483
Qy 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluLeuSerValMetThr 180
Db 484 AAGCTGCGAGCTGCCACACAGCTACATCTGCTGTGCATTGAGATAGCTTCATGACC 543
Qy 181 SerPheSerLys 184
Db 544 TCTTCTCCAAA 555

RESULT 13

US-10-036-869-37
; Sequence 37, Application US/10036869
; Publication No. US20020151516A1
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA
; ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Connolly, Bove, Lodge, & Hutz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/036,869
; FILING DATE: 29-No. US20020151516A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,526
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorris Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 565 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 37:
US-10-036-869-37
Alignment Scores:
Pred. No.: 7.78e-93 Length: 565
Score: 831.00 Matches: 154

Percent Similarity: 92.93% Conservative: 17
Best Local Similarity: 83.70% Mismatches: 13
Query Match: 86.02% Indels: 0
DB: 13 Gaps: 0
US-09-938-391-4 (1-184) x US-10-036-869-37 (1-565)
Qy 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
Db 11 CATACTCATCAGACTTTTTCAGCCAGTGCTCCACCTGGTGACACTGAACACCCCCCTGTCT 70
Qy 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaAla 40
Db 71 GAGGAGCATCGTGGTATCCGTGGAGCAGATTCCAGTGTCTCCAGCAAGCCGAGCGGTG 130
Qy 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 131 GGGCTGTCCGGCACCTTCCGGGGCTTTCCTGTCTCTTAGGCTGCAGGATCTCTATAGCATC 190
Qy 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
Db 191 GTGCGCCGTGCTGACCGGGGCTGTGGCCCATCGTCAACCTGAAGACGAGGTGCTATCT 250
Qy 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnGlyGlnLeuLysProGlyAlaArgIle 100
Db 251 CCAGCTGGGACTCCCTGTTTCTGGCTCCAGGCTCACTGCAACCCGGGGCCCGCATC 310
Qy 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 311 TTTTCTTTTGACGGCAGAGATGCTCTGAGACACCCAGCTGGCGCAGAGAGCGTATGG 370
Qy 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
Db 371 CACGGCTCGAGCCCGAGTGGCGGAGGCTGATGGAGAGTTACTGTGAGACATGGCGAAT 430
Qy 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
Db 431 GAACTACTGGGGCTACAGTCAAGGCTCTCCCTGCTGTGAGGAGGCTCTCTGGAAACAG 490
Qy 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluLeuSerValMetThr 180
Db 491 AAAGCTGCGAGCTGCCAACACAGCTACATCGTCTGTCATTGAGATAGCTTCATGACC 550
Qy 181 SerPheSerLys 184
Db 551 TCTTCTCCAAA 562
RESULT 14
US-09-998-831-12
; Sequence 12, Application US/09998831
; Patent No. US20020119153A1
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY
; TITLE OF INVENTION: INHIBITING VEGF
; FILE REFERENCE: 4001.002584
; CURRENT APPLICATION NUMBER: US/09/998,831
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 09/561,108
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: OLIGONUCLEOTIDE
; NAME/KEY: CDS
; LOCATION: (1)..(573)
US-09-998-831-12

Alignment Scores:
Pred. No.: 7,92e-93 Length: 573
Score: 831.00 Matches: 154
Percent Similarity: 92.93% Conservative: 17
Best Local Similarity: 83.70% Mismatches: 13
Query Match: 86.02% Indels: 0
DB: 10 Gaps: 0

US-09-938-391-4 (1-184) x US-09-998-831-12 (1-573)

QY 1 HisThrHisGlnAspPheGlnLeuValHisLeuValAlaLeuAsnSerProGlnPro 20
|||
DB 22 CATACTCATCAGGACTTTTCAGCCAGTCTCCACCTGGTGCGACCTGAACACCCCTGTCT 81
|||
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaA 40
|||
DB 82 GGAGGCATCGGTGTATCTCGTGAGCAGATTTCCAGTCTTCAGCAACCCGAGCCGTG 141
|||
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTySerIle 60
|||
DB 142 GGCTGTCCGGGACCTTCCGGGCTTTCTGTCTCTAGGCTGCAGGATCTCTATAGCATC 201
|||
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
|||
DB 202 GTGCGCGTGTGACCGGGGTCTGTGCCCATCTCAACCTGAAGCAGCAGGTGTCTATCT 261
|||
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuTySerProGlyAlaArgIle 100
|||
DB 262 CCAGCTGGGACTCCCTGTCTTCTGGCTCCCGGGTCAACTGCAACCCCGGGCCCGCATC 321
|||
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
|||
DB 322 TTTTCTTTTGACGGCAGAGATGCTCTGACACCCAGCTGCGCCGAGAGAGCGTATGG 381
|||
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyCysGluThrTrpArgThr 140
|||
DB 382 CACGGCTCGACCCAGTGGCGGAGGCTGTGTGAGAGTACTGTGAGACATGCGCAACT 441
|||
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuAlaGlyArgLeuGluGln 160
|||
DB 442 GAACTACTGGGGCTACAGCTAGGCTCTCTCCCTGTGTGAGCAGGCTCTCTGGAAACAG 501
|||
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
|||
DB 502 AAAGCTGCGAGCTGCCACACAGCTACATCGTCTGTGATTGAGATAGCTTCATGACC 561
|||
QY 181 SerPheSerLys 184
|||
DB 562 TCTTTCTCCAAA 573
|||

RESULT 15
US-10-292-418-17
; Sequence 17, Application US/10292418
; Publication No. US20030139365A1
; GENERAL INFORMATION:
; APPLICANT: Lo, Kin-Ming
; APPLICANT: Li, Yue
; APPLICANT: Gillies, Stephen D
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
; TITLE OF INVENTION: Immunofusins
; FILE REFERENCE: LEX-006C1
; CURRENT APPLICATION NUMBER: US/10/292,418
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/383,315
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: US 60/097,883
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 552
; TYPE: DNA

ORGANISM: Mus musculus
FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(552)
; OTHER INFORMATION: endostatin
US-10-292-418-17

Alignment Scores:
Pred. No.: 1.76e-92 Length: 552
Score: 828.00 Matches: 153
Percent Similarity: 92.93% Conservative: 18
Best Local Similarity: 83.11% Mismatches: 13
Query Match: 85.71% Indels: 0
DB: 12 Gaps: 0

US-09-938-391-4 (1-184) x US-10-292-418-17 (1-552)

QY 1 HisThrHisGlnAspPheGlnLeuValHisLeuValAlaLeuAsnSerProGlnPro 20
|||
DB 1 CATACTCATCAGGACTTTTCAGCCAGTCTCCACCTGGTGCGACCTGAACACCCCTGTCT 60
|||
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaA 40
|||
DB 61 GGAGGCATCGGTGTATCTCGTGAGCAGATTTCCAGTCTTCAGCAACCCGAGCCGTG 120
|||
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTySerIle 60
|||
DB 121 GGCTGTCCGGGACCTTCCGGGCTTTCTGTCTCTAGGCTGCAGGATCTCTATAGCATC 180
|||
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
|||
DB 181 GTGCGCGTGTGACCGGGGTCTGTGCCCATCTCAACCTGAAGCAGCAGGTGTCTATCT 240
|||
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuTySerProGlyAlaArgIle 100
|||
DB 241 CCAGCTGGGACTCCCTGTCTTCTGGCTCCCGGGTCAAGTCAACCCCGGGCCCGCATC 300
|||
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
|||
DB 301 TTTTCTTTTGACGGCAGAGATGCTCTGACACCCAGCTGCGCCGAGAGAGCGTATGG 360
|||
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyCysGluThrTrpArgThr 140
|||
DB 361 CACGGCTCGACCCAGTGGCGGAGGCTGTGTGAGAGTACTGTGAGACATGCGCAACT 420
|||
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuAlaGlyArgLeuGluGln 160
|||
DB 421 GAACTACTGGGGCTACAGGTACAGCTCAGGCTCTCTCCCTGTGTGAGCAGGCTCTCTGGAAACAG 480
|||
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
|||
DB 481 AAAGCTGCGAGCTGCCACACAGCTACATCGTCTGTGATTGAGATAGCTTCATGACC 540
|||
QY 181 SerPheSerLys 184
|||
DB 541 TCTTTCTCCAAA 552
|||

Search completed: August 18, 2003, 00:40:24
Job time : 224.556 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 17, 2003, 19:46:50 ; Search time 1630.67 Seconds
(without alignments)
2742.451 Million cell updates/sec

Title: US-09-938-391-4
Perfect score: 966
Sequence: 1 HTHQDFQLVHLVALNSQP.....CRHAFVVLCIENSVMTSFSK 184

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=slp
-Q=/cgn2_1/USPTO.spool_p/US09938391/runat_04082003_130656_27102/app.query.fasta_1.718
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938391@cgn_1.1.4399 @runat_04082003_130656_27102 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	835	86.4	707	10	BE908201	BE908201 601500458
2	835	86.4	757	10	BE908253	BE908253 601502237
3	835	86.4	881	14	CD105862	CD105862 AGENCOURT
4	831	86.0	843	10	BF385854	BF385854 602046021
5	831	86.0	874	12	BI412588	BI412588 602990468
6	823	85.2	715	14	CB596713	CB596713 AGENCOURT
7	822	85.1	551	10	BF074459	BF074459 221883 MA
8	818	84.7	611	10	AW911243	AW911243 ur83h10.y
9	806	83.4	753	12	BI904605	BI904605 603168411
10	805	83.3	897	12	BI080524	BI080524 602877005
11	803	83.1	782	9	AI326391	AI326391 mm18f09.x
12	794	82.2	733	10	EG072504	EG072504 H311D11-
13	793	82.1	723	12	BI247582	BI247582 602860041
14	784	81.2	944	13	BUS59398	BUS59398 AGENCOURT
15	781	80.8	915	10	BF166139	BF166139 601776586
16	775.5	80.3	929	13	BQ672290	BQ672290 AGENCOURT
17	774.5	80.2	835	12	BI526580	BI526580 602925454
18	772	79.9	720	12	BI147444	BI147444 602814008
19	770	79.7	703	13	BU615520	BU615520 UI-H-FGO-
20	763.5	79.0	979	13	BQ673186	BQ673186 AGENCOURT
21	753	78.0	683	12	BM683067	BM683067 UI-E-E01-
22	752	77.8	690	12	BI219399	BI219399 602936756
23	751	77.7	587	9	AA288198	AA288198 vbl5b10.r
24	750	77.6	795	14	CB601615	CB601615 AGENCOURT
25	744	77.0	652	13	BU352506	BU352506 603527982
26	744	77.0	657	14	CB444165	CB444165 695295 MA
27	739	76.5	682	9	AW089583	AW089583 xg20f03.x
28	739	76.5	832	10	BG387051	BG387051 602454749
29	735	76.1	884	12	BI161007	BI161007 602865213
30	732	75.8	1093	13	BQ723254	BQ723254 AGENCOURT
31	731	75.7	614	13	BU459935	BU459935 603367327
32	731	75.7	706	13	BU439577	BU439577 603208252
33	726	75.2	592	13	BQ567187	BQ567187 gi80h10.y
34	726	75.2	618	9	AV696242	AV696242 AV696242
35	726	75.2	633	11	AK014292	AK014292 Mus muscu
36	726	75.2	653	10	BF384828	BF384828 602046324
37	723.5	74.9	772	10	BF178591	BF178591 601807962
38	722	74.7	947	13	BUS56872	BUS56872 AGENCOURT
39	718	74.3	715	9	AU125614	AU125614 AU125614
40	707	73.2	675	10	BF385574	BF385574 602045824
41	705	73.0	661	13	BQ396997	BQ396997 NISC pg25
42	699	72.4	650	9	AI858615	AI858615 w140f01.x
43	695	71.9	747	12	BG967333	BG967333 602833649
44	694	71.8	536	9	AA277488	AA277488 v8a2907.r
45	694	71.8	664	13	BU632049	BU632049 UI-H-FE1-

ALIGNMENTS

RESULT 1
BE908201
LOCUS BE908201 707 bp mRNA linear EST 20-OCT-2000
DEFINITION 601500458F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3902175 5', mRNA sequence.
ACCESSION BE908201
VERSION BE908201.1 GI:10402537
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 707)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLM9704 row: d column: 16
 High quality sequence stop: 688.

FEATURES
 source
 1. .707
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3902175"
 /tissue_type="epithelioid carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_70"
 /notes="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.1 kb. Library constructed by Life
 Technologies."
 BASE COUNT 113 a 250 c 225 g 119 t

Alignment Scores:
 Pred. No.: 1.1e-71 Length: 707
 Score: 835.00 Matches: 154
 Percent Similarity: 92.82% Conservative: 14
 Best Local Similarity: 85.08% Mismatches: 13
 Query Match: 86.44% Indels: 0
 DB: 10, Gaps: 0

US-09-938-391-4 (1-184) x BE908201 (1-707)

QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
 DB 11 CACAGCCACCGGACCTCCAGCGGTGCTCCACCTGCTGCGCTCAACAGCCCGCTGTCA 70
 QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaA 40
 DB 71 GCGCGCATCGGGGCACTCCCGGGCGGACTTCCAGTGTTCAGCAGCGCGGCGCGTG 130
 QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
 DB 131 GGGCTGGGGGACCTTCCGGGCTTCTGCTGCTGCGGCTGCGGACCTGTACAGCATC 190
 QY 61 ValArgAlaAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
 DB 191 GTCCGCGCTGCCAGCCGCGCAGCGGTGCTCCATCGTCAACCTCAAGCAGAGCTGTGTT 250
 QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 100
 DB 251 CCCAGTGGGAGGCTGTGTTCTCAGGCTCTGAGGGTCCGCTGAAGCCCGGGGACGCATC 310
 QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
 DB 311 TTCTCTTTGACGGCAAGGACGTCTTGAGGACACCCACCTGGGCCCCAGAGAGCGTGTG 370
 QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
 DB 371 CATGGCTCGGACCCCAACGGCGCGAGCTGACCGAGAGCTACTGTGAGACGTGCGGACG 430
 QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
 DB 431 GAGGCTCCCTCGGCCACGGGCCAGGGCTCTCTGCTGCTGGGGGCGAGGCTCTCGGGGCA 490
 QY 161 GluAlaAlaSerCysArgHisAlaPheValValCysIleGluAsnSerValMetThr 180

Db 491 AGTGGCGGAGGTGCCATCAGCCCTACATCGTCTGCTGATTGAGAACAGCTTCATGACT 550
 QY 181 Ser 181
 Db 551 GCC 553

RESULT 2
 BE906253
 LOCUS
 DEFINITION 757 bp mRNA linear EST 20-OCT-2000
 601502237F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3904208 5',
 mRNA sequence.
 ACCESSION BE906253
 VERSION BE906253.1 GI:10399595
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 757)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLM9709 row: i column: 09
 High quality sequence stop: 757.

FEATURES
 source

1. .757
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3904208"
 /tissue_type="epithelioid carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_70"
 /note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.1 kb. Library constructed by Life
 Technologies."

BASE COUNT 145 a 258 c 235 g 119 t

Alignment Scores:
 Pred. No.: 1.2e-71 Length: 757
 Score: 835.00 Matches: 154
 Percent Similarity: 92.82% Conservative: 14
 Best Local Similarity: 85.08% Mismatches: 13
 Query Match: 86.44% Indels: 0
 DB: 10, Gaps: 0

US-09-938-391-4 (1-184) x BE906253 (1-757)

QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
 DB 36 CACAGCCACCGGACCTTCAGCGGTGCTCCACCTGTTGCGCTCAACAGCCCGCTGTCA 95
 QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaA 40
 DB 96 GCGCGCATCGGGGCACTCCCGGGCGGACTTCCAGTGTTCAGCAGCGCGGCGCGTG 155
 QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
 DB 156 GGGCTGGGGGACCTTCCGGGCTTCTGCTGCTGCGGCTCAGACCTGTACAGCATC 215
 QY 61 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80


```

Plate: LLAM9531 row: i column: 05
High quality sequence stop: 761.
Location/Qualifiers
1..843
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4195660"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_L19"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 167 a 278 c 228 g 170 t
ORIGIN

Alignment Scores:
Pred. No.: 3,42e-71 Length: 843
Score: 831.00 Matches: 154
Percent Similarity: 92.93% Conservative: 17
Best Local Similarity: 83.70% Mismatches: 13
Query Match: 86.02% Indels: 0
DB: 10 Gaps: 0

US-09-938-391-4 (1-184) x BF385854 (1-843)
QY 1 HistHrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
DB 171 CATACTCATCAGGACTTTCAGCCAGTGTCTCCACCTGGTGGCACTGAACACCCCTGTCT 230
QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnAlaArgAlaA 40
DB 231 GGAGGATGCTGGTGTATCCGTGGAGCAGATTTCCAGTGTCTTCAGCAAGCCGAGCCG 290
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSer 60
DB 291 GGCTGTGGGCACCTTCGGGGCTTTCCTGTCTCTAGGCTGCAGATCTCTATAGCATC 350
QY 61 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 351 GTGCGCGCTGCTGACCGGGGCTGTGTGCCATCGTCAACCTGAAGACGAGGTGTATCT 410
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgGlie 100
DB 411 CCAGCTGGGACCTCCCTGTTTCTGGCTCCAGGGTCAACTGCAACCCGGGGCCGCATC 470
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
DB 471 TTTTCTTTTGACGGCAGAGATGCTCTGAGACACCCAGCTGGCGCGCAGAGAGCGTATGG 530
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
DB 531 CACGGCTCGGACCCACCTGCGGGGAGCTGTATGATGAGATGTACTGTGAGACATGCGCAACT 590
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
DB 591 GAAACTACTGGGCTACAGGTACGGCTCTCTCTGTCGAGCAGGCTCTGGAACAG 650
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
DB 651 AAAGTCGAGCTGCCACACAGCTACATCGTCTCTGTGCATTGAGATAGCTTCATGACC 710
QY 181 SerPheSerLys 184
DB 711 TCTTTCTCCAAA 722

RESULT 5
BI412588
LOCUS 874 bp mRNA linear EST 14-AUG-2001
DEFINITION 602990468F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5146409 5', mRNA sequence.

```

ACCESSION BI412588
 VERSION BI412588.1 GI:15173511
 KEYWORDS EST
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE 1 (bases 1 to 874)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1361 row: g column: 18
 High quality sequence start: 23
 High quality sequence stop: 808.
 Location/Qualifiers
 1..874
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CZECH II"
 /db_xref="taxon:10090"
 /clone="IMAGE:5146409"
 /tissue_type="pooled lung tumors"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NCI_CGAP_Lu33"
 /note="Organ: lung; Vector: pTT73D-Pac (Pharmacia) with a modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCTCTGTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 189 a 275 c 228 g 181 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.6e-71 Length: 874
 Score: 831.00 Matches: 154
 Percent Similarity: 92.93% Conservative: 17
 Best Local Similarity: 83.70% Mismatches: 13
 Query Match: 86.02% Indels: 0
 DB: 12 Gaps: 0
 US-09-938-391-4 (1-184) x BI412588 (1-874)
 QY 1 HistHrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
 DB 143 CATACTCATCAGGACTTTCAGCCAGTGTCTCCACCTGGTGGCACTGAACACCCCTGTCT 202
 QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnAlaArgAlaA 40
 DB 203 GGAGGATGCTGGTGTATCCGTGGAGCAGATTTCCAGTGTCTTCAGCAAGCCGAGCCG 262
 QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSer 60
 DB 263 GGGCTGTGGGACCTTCCGGGCTTTCCTGTCTTAGGCTGCAGATCTCTATAGCATC 322
 QY 61 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
 DB 323 GTGCGCGGTGTGACCGGGGGTCTGTGCCATCTGTCACCTGAAGACGAGGTGTATCT 382

```

QY 81 ProSerTprGluAlaLeuPheSerGlySerGluGlnLeuLysProGlyAlaArgile 100
Db 383 CCAGCTGGGACTCCCTGTTCTGGCTCCAGGGTCAACTGCAACCCGGGGCCGCATC 442
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTprProArgLysSerValTpr 120
Db 443 TTTTCTTTTTCAGCGGACAGATGCTCTGAGACACCCAGCCCTGGCCGAGAGCGTATGG 502
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTprArgThr 140
Db 503 CACGGCTCGGACCCAGTGGCGAGCGTGTAGAGAGTACTGTGAGACATGGCGAAT 562
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
Db 563 GAACTACTGGGGCTACAGTCAAGGCTCTCTGCTGTGAGGAGGCTCTGGAACAG 622
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluHisSerValMetThr 180
Db 623 AAAGCTGCAGAGTGCACACAGTACATGCTGCTGTGATTGAGAATAGCTTCATGACC 682
QY 181 SerPheSerLys 184
Db 683 TCTTCTCCAAA 694

RESULT 6
CB596713
LOCUS
DEFINITION
AGENCOURT 12990486 NIH MGC_178 Mus musculus cDNA clone
IMAGE:30299651 5', mRNA sequence.
CB596713
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgi.nci.nih.gov/.
1 (bases 1 to 715)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM64 row: b column: 12
High quality sequence stop: 567.
FEATURES
source
1..715
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30299651"
/lab_host="DHI0B (11-phase-resistant)"
/clone_lib="NIH MGC 178"
/notes="Organ: lung and heart; Vector: pDNR-LIB; Site 1:
SfiI (ggccattatggcc); Site 2: SfiI (ggcgctcgcc); cDNA
made by oligo-dt priming and directionally cloned. 5' and
3' adaptors were used in cloning as follows:
5'-AACGATGTGTATCAAGCAGATGGAGTGGCATTCAGCGCGG-3' and
5'-ATTCTAGAGCCGAGCGCGGACATG-dt(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 0.5 kb
size fraction. Library created in the laboratory of M.
Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."
BASE COUNT 151 a 205 c 201 g 158 t
ORIGIN
Alignment Scores:

```

```

Pred. No.: 1.66e-70 Length: 715
Score: 823.00 Matches: 153
Percent Similarity: 92.90% Conservative: 17
Best Local Similarity: 83.61% Mismatches: 13
Query Match: 85.20% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-4 (1-184) x CB596713 (1-715)
QY 2 ThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnProGly 21
Db 5 ACTCATCAGGACTTTTCAGCCAGTGTCCACCTGTGGGCACTGACACCCCTGTCTGGA 64
QY 22 GlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAlaGly 41
Db 65 GCATCGGTGGTATCCGTGGAGCAGATTTCAGTGTCTCCAGCAAGCCCGGCGGG 124
QY 42 LeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIleVal 61
Db 125 CTGTCGGGCACCTTCGGGCTTTCTCTGCTCTAGGCTGAGGATCTCTATAGCATCGTG 184
QY 62 ArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhePro 81
Db 185 CCCTGTCTGACCCGGGGTCTGTGCCATCTGTCACCTGAGGACGAGGTGCTATCTCCC 244
QY 82 SerTprGluAlaLeuPheSerGlySerGluGlnLeuLysProGlyAlaArgilePhe 101
Db 245 AGCTGGGACTCCCTGTTTCTGGCTCCAGGCTCAACTGCAACCCGGGGCCGCATCTTT 304
QY 102 SerPheAspGlyArgAspValLeuGlnHisProAlaTprProArgLysSerValTprHis 121
Db 305 TCTTTTGACGGCAGAGATGCTCTGAGACACCCAGCCTGGCCGAGAGAGCGTATGGCAC 364
QY 122 GlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTprArgThrGlu 141
Db 365 GGCTCGGACCCAGTGGCGGAGCGGTGTGAGAGAGTTACTGTGAGACATGGCGAATGAA 424
QY 142 AlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGlnGlu 161
Db 425 ACTACTGGGGCTACAGGTCAAGCCCTCTCTCTGCTGTCTGAGGAGGCTCTCTGGAACAGAA 484
QY 162 AlaAlaSerCysArgHisAlaPheValValLeuCysIleGluHisSerValMetThrSer 181
Db 485 GCTGCGAGCTGCCAACACAGCTACATGCTCTGTCATTGAGAATAGCTTCATGACCTCT 544
QY 182 PheSerLys 184
Db 545 TTCTCCAAA 553

RESULT 7
BF074459 551 bp mRNA linear EST 25-APR-2001
LOCUS
DEFINITION
BF074459 Bos taurus cDNA 5', mRNA sequence.
ACCESSION
VERSION
BF074459.1 GI:10867970
KEYWORDS
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 551)
Smith, T.P.L., Grosse, W.M., Preking, B.A., Roberts, A.J., Stone, R.T.,
Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett
, G.L., Heaton, M.P., Laegreid, W., Rohrer, G.A., Chitko-McKown, C.G.,
Perrea, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and
Keele, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
JOURNAL
MEDLINE 21180013
PUBMED 11282978
COMMENT
Contact: Smith TPL

```

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@engr.marc.usda.gov
Single pass sequencing. Bases called and alt. trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.

PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCACGACG
Plate: 81 row: F column: 9
Seq primer: ATTAGTGACACTAG.

FEATURES

Location/Qualifiers
1..551
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH108"
/clone_lib="WARC 2BOV"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, thymus,
semitendinosus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."

BASE COUNT 80 a 200 c 179 g 92 t

ORIGIN

Alignment Scores:
Pred. No.: 1,47e-70 Length: 551
Score: 822.00 Matches: 157
Percent Similarity: 94.32% Conservative: 9
Best Local Similarity: 89.20% Mismatches: 10
Query Match: 85.09% Indels: 0
DB: 10 Gaps: 0

US-09-938-391-4 (1-184) x BF074459 (1-551)

Qy 9 ValLeuHisLeuValAlaLeuAsnSerProGlnProGlyGlyMetArgGlyIleArgGly 28
Db 1 GTCTGACCTTGTGGCGCTCAACAGCCCGCATGCGGCGGCTGCGCGCATCCCGGC 60
Qy 29 AlaAspPheGlnCysPheGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAla 48
Db 61 GCCGACTTTCAGTGCTTCCAGCAGGCGCGCGCGGCTGCGCGCACCTTCCCGCG 120
Qy 49 PheLeuSerSerArgLeuGlnAspLeuTySerIleValArgAlaAspArgThrGly 68
Db 121 TTCCTGTCTCTCGGGTTCAGAGACCTGTACAGCATCGTGGCGCGCGCCGCGCAC 180
Qy 69 ValProValValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSer 88
Db 181 CTCCCGCTGCTCAACCTCAGGAGAGGTGCTGTTCTAGCTGGGAGGCTTGTCTCA 240
Qy 89 GlySerGluGlyGlnLeuLysProGlyValaArgIlePheSerPheAspGlyArgAspVal 108
Db 241 GGCTCCGAGGCCAGCTGAGCCGCGCGCGCATCTTCTCTCCAGCGGCAGATGTC 300
Qy 109 LeuGlnHisProAlaTrpProArgLysSerValTrpHisGlySerAspProSerGlyArg 128
Db 301 CTTTCAGCATCCACCTCGGCCAGAGAGAGGTGTCGACGAGTATCATCCCGCGCGC 360
Qy 129 ArgLeuThrAspSerTyTyCysGluThrTrpArgThrCluAlaProAlaThrGlyGln 148
Db 361 CGCTGACCGAGAGCTACTCGGAGACGTGGCGGACGAGCGGCGGCCAGCGCCAG 420
Qy 149 AlaSerSerLeuLeuAlaGlyArgLeuLeuGluGlnAlaAlaSerCysArgHisAla 168
Db 421 GCCTCTCTCGTGTGGCGGCGCGGCTGCTGAGCAGAAAGCCCGCGGTGCCAACGCC 480
Qy 169 PheValLeuLeuCysIleGluAsnSerValMetThrSerPheSerIys 184
Db 481 TTCAATTGCTCTGTCATCGAGAACAGCTTTCATGACCTCTCTCTCTCTCTCTCTCA 528

RESULT 8

AW911243

LOCUS

DEFINITION

ur83h10.y1 NCI CGAP Mam6 Mus musculus cDNA clone IMAGE:3156931 5'

similar to gb:DI7546 Mouse mRNA for collagen (MOUSE);, mRNA

sequence.

AW911243

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished

Other ESTs: ur83h10.x1

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey Green M.D.

cDNA library Preparation: Life Technologies, Inc.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

image.llnl.gov/image/html/resources.shtml

MGI:1059687

Seq primer: -40RP from Gibco

High quality sequence stop: 403.

Location/Qualifiers

1..611

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:3156931"

/sex="female, virgin"

/tissue_type="infiltrating ductal carcinoma"

/dev_stage="5 months"

/lab_host="DH108"

/clone_lib="NCI CGAP Mam6"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies. Investigator

providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 113 a 199 c 166 g 133 t

ORIGIN

Alignment Scores:

Pred. No.: 4.15e-70 Length: 611

Score: 818.00 Matches: 152

Percent Similarity: 91.26% Conservative: 15

Best Local Similarity: 83.06% Mismatches: 16

Query Match: 84.68% Indels: 0

DB: 10 Gaps: 0

US-09-938-391-4 (1-184) x AW911243 (1-611)

Qy 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20

Db 61 CATACTCATCAGACCTTTCAGCAGCTTCCACCTGCTGCTGCTGCTGCTGCTGCT 120

Qy 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAla 40

Db 121 GGAGGCATCGTGTATCGTGGAGCAGATTTCCAGTGTCTTCAGCAAGCCGAGCCGTG 180

Qy 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTySerIle 60

Db 181 GGGCTGTGGGACCTTCGGGCTTTCCTCTCTAGGCTGAGGATCTCTATAGCATC 240

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIle 60
 DB 592 GGCGTCTGGGACACCTTCGGGGCTTCTCTAGGCTGCAGGATCTCTATAGCATC 533

QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
 DB 532 GTGCGCGCTGCTGACCGGGGGTCTGTGCCATCTCAACCTGAAGACGAGGTGTATCT 473

QY 81 ProSerTrpGluAlaLeuPheSerGlySerGluGlnLeuLysProGlyAlaArgIle 100
 DB 472 CCCAGTGGGACTCCTCTGTTCTGCTCNCAGGCTCAACTGCAACCCGGGCCCCGCATC 413

QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
 DB 412 TTTTCTTTTTCAGCGCAGATGCTCTGAGACACCCAGCCTGGCGCGCAGAAGCGATGCG 353

QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
 DB 352 CACGGCTCGGACCCAGTGGCGGAGGCTGATGAGAGTTACTGTGAGACATGGCGAATC 293

QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuAlaGlyArgLeuLeuGluGln 160
 DB 292 GAAACTACTGGGGCTACAGTACGGCTCTCTCTGCTGTCAGCAGGCTCTCTGGAACAG 233

QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
 DB 232 AAAGCTGGAGCTGCCACACACTACATCTCTCTGCTGATGAGATAGCTTCATGACC 173

QY 181 SerPheSerLys 184
 DB 172 TCTTTCTCCAAA 161

RESULT 12
 BG072504/c

LOCUS H3111D11-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
 DEFINITION H3111D11 3', mRNA sequence.

ACCESSION BG072504
 VERSION BG072504.1 GI:12555073
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 733)
 Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac
 ,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H.
 III, Becker,K.G. and Ko,M.S.H.
 Genome-wide expression profiling of mid-gestation placenta and
 embryo using a 15,000 mouse developmental cDNA microarray
 Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
 20381348
 10922068

JOURNAL National Institute on Aging/National Institutes of Health
 MEDLINE 333 Caswell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 PUBMED Email: cdna@lgsun.grc.nia.nih.gov
 COMMENT This clone set has been freely distributed to the community. Please
 visit <http://lgsun.grc.nia.nih.gov/cdna/15k.html> for details.
 Plate: H3111 row: D column: 11
 Seq primer: -21M13 Forward
 High quality sequence stop: 733
 POLYA=Yes.

FEATURES
 source Location/Qualifiers
 1..733
 /orgasm="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="niaEST:H3111D11-3"
 /db_xref="taxon:10090"

/clone="H3111D11"
 /sex="Clones arrayed from a variety of cDNA libraries"
 /dev_stage="Clones arrayed from a variety of cDNA libraries"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse 15K cDNA Clone Set"
 /note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E42.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 Kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U.S.A. 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

BASE COUNT 155 a 202 c 204 g 172 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,18e-67 Length: 733
 Score: 794.00 Matches: 152
 Percent Similarity: 92.35% Conservative: 17
 Best Local Similarity: 83.06% Mismatches: 14
 Query Match: 82.19% Indels: 1
 DB: 10 Gaps: 0

US-09-938-391-4 (1-184) x BG072504 (1-733)

QY 2 ThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnProGly 21
 DB 731 ACTCATCAGGACTTTCAGCCAGTGTCTCCACCTGGTGGCACTGACACCCCTGTCTGGA 672

QY 22 GlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaGly 41
 DB 671 GGCATGCGTGGTATCGTGCAGCAGATTTCAGTGTTC-CAGCAAGCCCGAGCGTGGGG 613

QY 42 LeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIleVal 61
 DB 612 CTGTCGGGCACCTTCGGGGCTTCTCTGCTCTAGGCTCAGGATCTCTATAGCATCGTG 553

QY 62 ArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhePro 81
 DB 552 CGCCGTGCTGACCGGGGGTCTGTGCCATCTGTCACCTGAGGACGAGGTCTATCTCCC 493

QY 82 SerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIlePhe 101
 DB 492 AGCTGGGACTCCTCTGTTTCTGGCTCCAGGCTCAACTGCAACCCGGGCCCCGCATCTTT 433

QY 102 SerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrpHis 121
 DB 432 TCTTTTTCAGCGGAGAGATGCTCTGAGACACCCAGCCTGGCGCGCAGAAGACGCTATGGCAC 373

QY 122 GlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThrGlu 141
 DB 372 GGTCTGGACCCCGAGTGGCGGAGGCTGATGGAGAGTTACTGTGAGACATGGCGAATGAA 313

QY 142 AlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGluGln 161
 DB 312 ACTACTGGGGCTACAGTTCAGGCTCTCTCTCTCTCAGGACGAGTCTCTGGAACAGAA 253

QY 162 AlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThrSer 181
 DB 252 GCTGCGAGCTGCCAACACGCTACATCGTCTCTGTGATTTGAGAATAGCTTTCATGACCTCT 193


```

QY      182 PheSerIys 184
DB      192 TTCTCCAAA 184

RESULT 13
LOCUS   BI247582
DEFINITION 723 bp mRNA linear EST 17-JUL-2001
602960041F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:5125523 5',
mRNA sequence.
ACCESSION BI247582
VERSION   BI247582.1 GI:14792652
KEYWORDS  EST.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 723)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11307 row: a column: 12
High quality sequence stop: 665.
Location/Qualifiers
1. 723
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5125523"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NCI_CGAP_Li9"
/notes="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 158 a 209 c 201 g 155 t
ORIGIN
source
Alignment Scores:
Pred. No.: 145e-67 Length: 723
Score: 793.00 Matches: 148
Percent Similarity: 92.70% Conservative: 17
Best Local Similarity: 83.15% Mismatches: 13
Query Match: 82.09% Indels: 0
DB: 12 Gaps: 0

US-09-938-391-4 (1-184) x BI247582 (1-723)

QY      7 GlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnProGlyMetArgGlyIle 26
DB      1 CAGCCAGTGTCTCCACCTGGTGGCACTCAACACCCCTGCTGGAGGCATGCGTGTATC 60
QY      27 ArgGlyAlaAspPheGlnCysPheGlnAlaArgAlaAlaGlyLeuAlaGlyThrPhe 46
DB      61 CGTGGAGCAGATTTCAGTGTCTTCAGCAAGCCGAGCGGCTGTGGGGCTGTCCGGCACCTTC 120
QY      47 ArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIleValArgAlaAspArg 66
DB      121 CGGGCTTTCTCTCTCTAGCTCGAGATCTTATAGCATCTGGCCCGCTGTGACCGG 180
QY      67 ThrGlyValProValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeu 86
DB      181 GGGTCTGTGGCCCATCGTCAACCTGAAGGACGAGGTGCTATCTCCACGCTGGGACTCCCTG 240

```

```

QY      87 PheSerGlySerGluGlyGlnLeuIysProGlyValaArgGilePheSerPheAspGlyArg 106
DB      241 TTTTCTGGCTCCAGGGTCAACTGCACCCGGGGCCCACTTTTCTTTTGACGGCAGA 300
QY      107 AspValLeuGlnHisProAlaTrpProArgLysSerValTrpHisGlySerAspProSer 126
DB      301 GATGTCTTGAGACACCCAGCCTGGCCGAGAAAGCGGTATGGCAGCGCTCGGACCCAGT 360
QY      127 GlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThrGluAlaProAlaAlaThr 146
DB      361 GGGCGGAGGCTGATGGAGAGTTACTGTGAGACATGGCGAACTACTGGGGCTACA 420
QY      147 GlyGlnAlaSerSerLeuAlaGlyArgLeuLeuGlnGlnAlaAlaSerCysArg 166
DB      421 GGTACGGCTCTCTCTCTGTGAGCAGGCTCTGGAAACAGAAAGCTGGAGCTGCCAC 480
QY      167 HisAlaPheValValLeuCysIleGluAsnSerValMetThrSerPheSerIys 184
DB      481 AACAGCTACATCGTCTCTGTGCATTGAGAAATAGCTTCATGACCTCTTTCTCCAAA 534

RESULT 14
LOCUS   BU859398/c
DEFINITION 944 bp mRNA linear EST 16-OCT-2002
AGENCOURT 10434549 NIH MGC 109 Homo sapiens cDNA clone
IMAGE:6650260 5', mRNA sequence.
ACCESSION BU859398
VERSION   BU859398.1 GI:24044390
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 944)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM2895 row: d column: 04
High quality sequence stop: 572.
Location/Qualifiers
1. 944
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6650260"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_109"
/notes="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 163 a 331 c 260 g 190 t
ORIGIN
source
Alignment Scores:
Pred. No.: 158e-66 Length: 944
Score: 784.00 Matches: 144
Percent Similarity: 92.44% Conservative: 15
Best Local Similarity: 83.72% Mismatches: 13

```



```
Query Match: 81.16% Indels: 0
DB: 13 Gaps: 0
US-09-938-391-4 (1-184) x BU859398 (1-944)

QY 10 LeuHisLeuValAlaLeuAsnSerProGlnProGlyGlyMetArgGlyLeuArgGlyAla 29
DB 684 GTCCACATGTTGGCTGACAGCCGCCCTGTGTCAGCGCGCATCGGGGACATCGCGGGCCC 625
QY 30 AspPheGlnCysPheGlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPhe 49
DB 624 GACTTCCAGTGCTTCAGCAGCGCGCGCGCTGGGGCTGGCGGACCTTCCGGCCCTTC 565
QY 50 LeuSerSerArgLeuGlnAspLeuTyrsSerIleValArgArgAlaAspArgThrGlyVal 69
DB 564 CTGTCTCGCGCTGCAGGACCTGTACAGCATGTCGTGCGCGTGCAGCAGCGCGCGTG 505
QY 70 ProValValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGly 89
DB 504 CCCATCGTCAACCTCAAGGACGAGCTGCTGTTTCCAGCTGGGAGCTCTGTTCTCAGGC 445
QY 90 SerGluGlyGlnLeuLysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeu 109
DB 444 TCTGAGGGTCCGCTGAAGCCCGGGGACGATCTTCTCTTTCAGCGCAGGACGTCCTG 385
QY 110 GlnHisProAlaTrpProArgLysSerValTrpHisGlySerAspProSerGlyArgArg 129
DB 384 AGGCACCCACCTGGCCCCAGAGAGCGTGTGGCATGGTTCGACCCCAACGGGGCAGG 325
QY 130 LeuThrAspSerTyrcysGluThrTrpArgThrGluAlaProAlaAlaThrGlyGlnAla 149
DB 324 CTGACCGAGAGTACTGTGAGAGCTGGCGGACGAGGCTCCCTCGGCCACGCGGCAGGC 265
QY 150 SerSerLeuLeuAlaGlyArgLeuGluGlnGluAlaAlaSerCysArgHisAlaPhe 169
DB 264 TCTCTGCTCTGGGGGCGAGCTCTCGGGCAGAGTGGCGGAGTGCAGCTCACGCCCTAC 205
QY 170 ValValLeuCysIleGluAsnSerValMetThrSer 181
DB 204 ATCGTGCTCTGATTGAGAACAGCTTCATGACTGCC 169

RESULT 15
LOCUS BF166139
DEFINITION 601776586F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:4018317 5',
mRNA sequence.
ACCESSION BF166139
VERSION BF166139.1 GI:11046491
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9268 row: k column: 22
High quality sequence stop: 722.
Location/Qualifiers
1..915
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
```

```
/db_xref="taxon:10090"
/clone="IMAGE:4018317"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
/lab_host="VDH10B"
/clone_lib="NCI_CGAP_Lu29"
/Note="Organ: lung; Vector: pCMV-SPORT6; Site: 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 184 a 272 c 264 g 195 t
ORIGIN
```

Alignment Scores:

```
Pred. No.: 2,978-66 Length: 915
Score: 781.00 Matches: 153
Percent Similarity: 92.3% Conservative: 17
Best Local Similarity: 83.1% Mismatches: 14
Query Match: 80.8% Indels: 2
DB: 10 Gaps: 0
```

US-09-938-391-4 (1-184) x BF166139 (1-915)

```
QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
DB 21 CATACTCATCAGGACTTTCAGCCAGTGTCCACCTGGTGGCACTGAACACCCCTGTCT 80
QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
DB 81 GGAGCGATCGCGGTATCCGTGGAGCAGATTTCAGTGTCTCCAGCAAGCCGCGCGTG 140
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrsIle 60
DB 141 GGGCTGTGGGACCTTCCGGGCTTCTGTCTCTTAGGCTGCAGATCTCTATAGCATC 200
QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
DB 201 GTGCGCGTGTCTGACCGGGGTCTGTGCCATCGTCAACCTGAAGCAGAGGTGTCTAT 260
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgIle 100
DB 261 CCCAGTGGGACTCTCTGTTTCTGGTCCAGGGTCAATTGCAACCCCGGGCCGCGATC 320
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
DB 321 TTTTC-TTTGACGCGAGAGATGCTCTGACACACCCAGCTGGCCGAGAGCGTATGG 379
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrcysGluThrTrpArgThr 140
DB 380 CATGGCTCGGACCCAGTGGCGGAGGCTGTATGGAGAGTTACTGTGAGACATGGCGGACT 439
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
DB 440 GAAACTACTGGGGCTACAGGTGAGGCTCCCTCCCTGCTGTGAGGAGGCTCTCGGAACAG 499
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
DB 500 AAAGTGCAGCTGCAC-AACAGCTACATCTGCTGTGATTTGAGATAGTTCATGACC 558
QY 181 SerPheSerLys 184
DB 559 TCTTCTCCAAA 570
```

Search completed: August 17, 2003, 22:57:47
Job time : 1633.67 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 17, 2003, 23:00:16 ; Search time 2928.89 Seconds
(without alignments)
2570.043 Million cell updates/sec

Title: US-09-938-391-4
Perfect score: 184
Sequence: 1 HTHQDFQLVHLVALNSQP.....CRHAFVVLCIENSVMTSFSK 184

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Word size: 1

Total number of hits satisfying chosen parameters: 5773148

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n model -DEV=xl
-Q/cn2_1/USPTO_spool_p/US0938391/runat_04082003_130738_27585/app_query.fasta_1.718
-DB=GenEmbl -QFMT=fastap -SURFIX=p2noli.rge -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0938391 @CGN 1 1 5066 @runat_04082003_130738_27585 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1:	gb.ba.*	GenEmbl.*
2:	gb.htg.*	
3:	gb.in.*	
4:	gb.om.*	
5:	gb.ov.*	
6:	gb.pat.*	
7:	gb.ph.*	
8:	gb.pl.*	
9:	gb.pr.*	
10:	gb.ro.*	
11:	gb.sts.*	
12:	gb.sv.*	
13:	gb.un.*	
14:	gb.vi.*	
15:	em.ba.*	
16:	em.fun.*	
17:	em.hum.*	
18:	em.in.*	
19:	em.mu.*	
20:	em.om.*	
21:	em.or.*	
22:	em.ov.*	
23:	em.pat.*	
24:	em.ph.*	
25:	em.pl.*	
26:	em.ro.*	
27:	em.sts.*	
28:	em.un.*	

Result No.	Score	Match	Length	DB	ID	Description
1	184	100.0	555	6	AX399631	Sequence
2	184	100.0	829	6	AX399629	Sequence
3	27	14.7	5279	5	AF083440	Gallus ga
4	26	14.1	534	6	AR193166	Sequence
5	26	14.1	537	6	AX100092	Sequence
6	26	14.1	546	6	AR193165	Sequence
7	26	14.1	549	6	AX100086	Sequence
8	26	14.1	552	6	AX395662	Sequence
9	26	14.1	552	6	AX473835	Sequence
10	26	14.1	552	6	BD081407	Fused pro
11	26	14.1	555	9	AF184060	Homo sapi
12	26	14.1	564	9	AF416592	Homo sapi
13	26	14.1	650	6	AF370853	Homo sapi
14	26	14.1	786	9	AF282883	Homo sapi
15	26	14.1	900	6	AX370851	Sequence
16	26	14.1	2364	9	AK098216	Homo sapi
17	26	14.1	2637	9	BC033715	Homo sapi
18	26	14.1	3394	6	AX409531	Sequence
19	26	14.1	3394	6	IS1045	Sequence 4
20	26	14.1	3394	9	HUMCOL18AX	Human colla
21	26	14.1	5408	9	AF018082	Homo sapi
22	26	14.1	5929	9	AF018081	Homo sapi
23	26	14.1	131563	9	HSE12310	Homo sapi
24	26	14.1	223201	9	HS53110	Homo sapi
25	26	14.1	340000	9	HS21C102	Homo sapi
26	23	12.5	515	10	RNO236873	Rattus no
27	23	12.5	552	10	AF257775	Mus muscu
28	23	12.5	555	6	BD102800	Process f
29	23	12.5	555	6	BD128313	Endostati
30	23	12.5	558	6	AR268609	Sequence
31	23	12.5	565	6	AR100642	Sequence
32	23	12.5	565	6	E34073	Carrier/DNA
33	23	12.5	573	6	AR183471	Sequence
34	23	12.5	573	6	AR183520	Sequence
35	23	12.5	573	6	AR217279	Sequence
36	23	12.5	573	6	AR282751	Sequence
37	23	12.5	573	6	AX042272	Sequence
38	23	12.5	573	6	AX490666	Sequence
39	23	12.5	624	6	AX084406	Sequence
40	23	12.5	624	6	AX128409	Sequence
41	23	12.5	639	11	BV071218	Sequence
42	23	12.5	851	10	AF189709	Rattus no
43	23	12.5	1615	10	MMU03714	Mus musculu
44	23	12.5	3129	10	MMCOL18A11	Mus musculu
45	23	12.5	4031	6	IS1044	Sequence 1

ALIGNMENTS

RESULT 1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

AX399631
LOCUS AX399631 555 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 3 from Patent EP1191036.
ACCESSION AX399631
VERSION AX399631.1 GI:213335410
KEYWORDS
SOURCE
ORGANISM
Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1
REFERENCE
AUTHORS Sheppard, M.G. and Tong, X.
TITLE Methods and compositions for diagnosing and treating disorders involving angiogenesis
JOURNAL Patent: EP 1191036-A 3 27-MAR-2002;
Pfizer Products Inc. (US)
FEATURES
source Location/Qualifiers
1..555
/organism="Canis familiaris"
/mol_type="genomic DNA"
/db_xref="taxon:9615" 83 t
BASE COUNT 78 a 203 c 191 g 83 t
ORIGIN
Alignment Scores:
Pred. No.: 2,66e-179 Length: 555
Score: 184.00 Matches: 184
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-938-391-4 (1-184) x AX399631 (1-555)
QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
Db 1 CACACCCACCAGGACTTCCAGCTGGTGTGCTGCACCTGGTGGCCCTGAACAGCCGCGAGCG 60
QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaA 40
Db 61 GCGGCGATCGAGGCATCCGGGAGCGGACTTCCAGTGTCTCCAGCAGCGCGCGCGG 120
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 121 GGCGTGGCGGCACCTTCCGGGCTTCTCTGCTGCTGGCTGCAGGACCTCTACAGCATC 180
QY 61 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
Db 181 GTGCGCGCGCGCCAGCCGCGGCTGCGGCTGCTCAACCTCAGGAGCGAGGTGCTCTTC 240
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyAlaArgIle 100
Db 241 CCAGCTGGAGGACCTTATCTCGGGCTCCAGGGCCAGCTGAAGCCCGGGCCCGCATC 300
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 301 TTCTCTTCAGCGCAGAGATGCTCTGCAGACCCCGCTGGCCCGCGAGAGCGGTGG 360
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
Db 361 CACGGCTCCGACCCAGCGGGCCCGCTGACCGACAGCTACTTGCAGAGACGTGGCGGAGC 420
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
Db 421 GAGGCCCCGGCGCCACCGGGCAGGGCTGCTGCTGCTGGCGGGCAGGCTGTGGAGCAG 480
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
Db 481 GAGCGCGCGAGCTGCCGCCAGCGCTTCTGTTGCTCTCTGATCGAGAACAGCGTCATGACC 540
QY 181 SerPheSerLys 184
Db 541 TCCTTCTCCAAG 552

RESULT 2
AX399629
LOCUS AX399629 829 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 1 from Patent EP1191036.
ACCESSION AX399629
VERSION AX399629.1 GI:213335409
KEYWORDS
SOURCE
ORGANISM
Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1
REFERENCE
AUTHORS Sheppard, M.G. and Tong, X.
TITLE Methods and compositions for diagnosing and treating disorders involving angiogenesis
JOURNAL Patent: EP 1191036-A 1 27-MAR-2002;
Pfizer Products Inc. (US)
FEATURES
source Location/Qualifiers
1..829
/organism="Canis familiaris"
/mol_type="genomic DNA"
/db_xref="taxon:9615" 113 t
BASE COUNT 124 a 314 c 278 g 113 t
ORIGIN
Alignment Scores:
Pred. No.: 3,84e-179 Length: 829
Score: 184.00 Matches: 184
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-938-391-4 (1-184) x AX399629 (1-829)
QY 1 HisThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
Db 139 CACACCCACCAGGACTTCCAGCTGGTGTGCTGCACCTGGTGGCCCTGAACAGCCGCGAGCG 198
QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaA 40
Db 199 GCGCGCATCGAGGCATCCGGGAGCGGACTTCCAGTGTCTCCAGCAGCGCGCGCGG 258
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 259 GGCGTGGCGGCACCTTCCGGGCTTCTGCTGCTGGCTGCAGGACCTCTACAGCATC 318
QY 61 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
Db 319 GTGCGCGCGCGCCAGCCGCGGCTGCGGCTGCTCAACCTCAGGAGCGAGGTGCTCTTC 378
QY 81 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuLysProGlyAlaArgIle 100
Db 379 CCAGCTGGAGGACCTTATCTCGGGCTCCAGGGCCAGCTGAAGCCCGGGCCCGCATC 438
QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
Db 439 TTCTCTTCAGCGCAGAGATGCTCTGCAGACCCCGCTGGCCCGCGAGAGCGGTGG 498
QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
Db 499 CACGGCTCCGACCCCGCGGGCCCGCTGACCGACAGCTACTTGCAGAGACGTGGCGGAGC 558
QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuGluGln 160
Db 559 GAGGCCCCGGCGCCACCGGGCAGGGCTGCTGCTGCTGGCGGGCAGGCTGTGGAGCAG 618
QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
Db 619 GAGGCGCGAGCTGCGCCACGCCCTTCTGTTGCTCTGATCGAGAACAGCGTCATGACC 678
QY 181 SerPheSerLys 184

```
Db 679 TCCTCTCAAG 690
|||||
RESULT 3
AF083440 5279 bp mRNA linear VRT 31-AUG-2001
LOCUS Gallus gallus collagen XVIII mRNA, complete cds.
DEFINITION AF083440
ACCESSION AF083440
VERSION AF083440.2 GI:15383904
KEYWORDS Gallus gallus (chicken)
SOURCE Gallus gallus
ORGANISM Gallus gallus
REFERENCE 1 (bases 1 to 5279)
AUTHORS Halfter,W., Dong,S., Schurer,B. and Cole,G.J.
TITLE Collagen XVIII is a basement membrane heparan sulfate proteoglycan
JOURNAL J. Biol. Chem. 273 (39), 25404-25412 (1998)
MEDLINE 98411346
PUBMED 9738008
REFERENCE 2 (bases 1 to 5279)
AUTHORS Halfter,W. and Dong,S.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1998) Neurobiology, University of Pittsburgh,
3500 Terrace Street, Pittsburgh, PA 15261, USA
REFERENCE 3 (bases 1 to 5279)
AUTHORS Halfter,W. and Dong,S.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-2001) Neurobiology, University of Pittsburgh,
3500 Terrace Street, Pittsburgh, PA 15261, USA
REMARK Sequence update by submitter
COMMENT On Aug 31, 2001 this sequence version replaced gi:3493142.
FEATURES
source
1..5279
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
8..4042
/codon_start=1
/product="collagen XVIII"
/protein_id="AAC33294.2"
/db_xref="GI:15383905"
/translation="MRPGCAPRRLLILGLFVLLLPAAQSEPNLSLELIGDP
PPEELIKYVGNPNVGVYVPSANTQVARYHLPSFVDESLFHIOPTTFRAGMLF
AITDTSQSIYYGVKLSLDMQKQIIFYYTPGSSVAAATFTVPLLNQWTRPAI
RGDRAAEQEEEDDAEASGDFGSAEDRHHPGKDKGIPGLLDVAVFTSPVVEG
SGTRSSAGSQQAERTVREQLQVSTGTGPKGKEGKGERGPKGDSGTSLGTGAT
KGEKGEKBLGIKSGAGFGYVPGSKGQKGPBPGLRHTDMSLEQVTPGPGPT
GPGKDGAPRGDGEDKPKGEMGQPGFTGESGQKGEKGDVPGVGRGPRGP
PGPPGGRSSQDGLTFIDMGSGFGDLETLRPGPPGPPGPPGPPGPPGPPGPPG
GNMSTDLPGRLDGTGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
PGPKGSGAGAPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
LPATGSGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
PGQPGPKNRSGSGGKGGQGVGLPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
SHAGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
PGPMGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP
PGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP
SASLRGDKGKGEAGKGEKGGPSSTLYGPPGPPGPPGPPGPPGPPGPPGPPG
PGQPGGIGIKGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG
LRTMPTQALMSAAHLEGGGLIFLADROELVIRLGGFRVRLLEHTLVPSALDNE
VYDKLPSIHYGAQVQVPHLRNHPPTARPRGDEGVANQHHLPQPLLQOHELINS
YTIHRMFPAPVAHVHQDFQALHLVALNTLPSLGMRGIRGADFCQFOQAQVGLAG
TFRAFSLSLQDLSYLVREADRTAVPVLNRLDEVLFSNWEALFTSEAPLRAGARILS
FGKDTLQDSAWPQKSIWHSQDAKGRRLPESYCEAWRTDERTGTSQASSLSGKLEQ
SASSOCHAFVLCIENSFWTAAK"
8..85
sig_peptide
3'UTR 4043..5279
BASE COUNT 1057 a 1668 c 1631 g 923 t
ORIGIN
Alignment Scores:
```

```

FEATURES             Location/Qualifiers
     source            1..537
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
BASE COUNT           79 a 189 c 176 g 93 t
ORIGIN
Alignment Scores:
Pred. No.:            4,69e-17      Length:      537
Score:                26.00         Matches:    26
Percent Similarity:   100.00%       Conservative: 0
Best Local Similarity: 100.00%       Mismatches: 0
Query Match:         14.13%         Indels:     0
DB:                  6              Gaps:        0

US-09-938-391-4 (1-184) x AX100092 (1-537)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
   |||||
Db 109 GGGCTGGCGGGACCTTCGCGCTTCCTGCTCGCGCTCGAGGACCTGTACAGCATC 168
   |||||
QY 61 ValArgAlaAspArg 66
   |||||
Db 169 GTGCGCGGTGCGACCGC 186

RESULT 6
AR193165
LOCUS AR193165 546 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 4 from patent US 6346510.
ACCESSION AR193165
VERSION AR193165.1 GI:20239130
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS O'Reilly,M.S. and Folkman,M.Judah.
TITLE Therapeutic antiangiogenic endostatin compositions
JOURNAL Patent: US 6346510-A 4 12-FEB-2002;
FEATURES             Location/Qualifiers
     source            1..546
                        /organism="unknown"
BASE COUNT           80 a 196 c 177 g 93 t
ORIGIN

Alignment Scores:
Pred. No.:            4,76e-17      Length:      546
Score:                26.00         Matches:    26
Percent Similarity:   100.00%       Conservative: 0
Best Local Similarity: 100.00%       Mismatches: 0
Query Match:         14.13%         Indels:     0
DB:                  6              Gaps:        0

US-09-938-391-4 (1-184) x AR193165 (1-546)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
   |||||
Db 121 GGGCTGGCGGGACCTTCGCGCTTCCTGCTCGCGCTCGAGGACCTGTACAGCATC 180
   |||||
QY 61 ValArgAlaAspArg 66
   |||||
Db 181 GTGCGCGGTGCGACCGC 198

RESULT 7
AX100086
LOCUS AX100086 549 bp DNA linear PAT 02-APR-2001
DEFINITION Sequence 4 from Patent WO0119989.
ACCESSION AX100086
VERSION AX100086.1 GI:13539061
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE
AUTHORS Liang,H., Sim,K.L., Chang-Murad,A., Zhou,X., Madsen,J.,
          Boerner,R.J., Bermejo,L.L., Mistry,F.R., Shepard,S.R. and
          Schrimsher,J.L.
TITLE Method of producing and purifying endostatin?tm protein
JOURNAL Patent: WO 0119989-A 4 22-MAR-2001;
FEATURES             Location/Qualifiers
     source            1..549
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
BASE COUNT           82 a 196 c 178 g 93 t
ORIGIN

Alignment Scores:
Pred. No.:            4,78e-17      Length:      549
Score:                26.00         Matches:    26
Percent Similarity:   100.00%       Conservative: 0
Best Local Similarity: 100.00%       Mismatches: 0
Query Match:         14.13%         Indels:     0
DB:                  6              Gaps:        0

US-09-938-391-4 (1-184) x AX100086 (1-549)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
   |||||
Db 121 GGGCTGGCGGGACCTTCGCGCTTCCTGCTCGCGCTCGAGGACCTGTACAGCATC 180
   |||||
QY 61 ValArgAlaAspArg 66
   |||||
Db 181 GTGCGCGGTGCGACCGC 198

RESULT 8
AX395662
LOCUS AX395662 552 bp DNA linear PAT 18-MAY-2002
DEFINITION Sequence 30 from Patent WO0193897.
ACCESSION AX395662
VERSION AX395662.1 GI:21066471
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS Sim,K.L. and Macdonald,N.J.
TITLE Angiotatin and endostatin binding proteins and methods of use
JOURNAL Patent: WO 0193897-A 30 13-DEC-2001;
FEATURES             Location/Qualifiers
     source            1..552
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
BASE COUNT           83 a 196 c 179 g 94 t
ORIGIN

Alignment Scores:
Pred. No.:            4,81e-17      Length:      552
Score:                26.00         Matches:    26
Percent Similarity:   100.00%       Conservative: 0
Best Local Similarity: 100.00%       Mismatches: 0
Query Match:         14.13%         Indels:     0
DB:                  6              Gaps:        0

US-09-938-391-4 (1-184) x AX395662 (1-552)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
   |||||
Db 121 GGGCTGGCGGGACCTTCGCGCTTCCTGCTCGCGCTCGAGGACCTGTACAGCATC 180
   |||||

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

Li, H., Sim, K. L., Chang-Murad, A., Zhou, X., Madsen, J., Boerner, R. J., Bermejo, L. L., Mistry, F. R., Shepard, S. R. and Schrimsher, J. L.

Method of producing and purifying endostatin?tm protein

Patent: WO 0119989-A 4 22-MAR-2001;

EntreMed, Inc. (US)

Location/Qualifiers

source 1..549

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT 82 a 196 c 178 g 93 t

ORIGIN

Alignment Scores:

Pred. No.: 4,78e-17 Length: 549

Score: 26.00 Matches: 26

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 14.13% Indels: 0

DB: 6 Gaps: 0

US-09-938-391-4 (1-184) x AX100086 (1-549)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60

Db 121 GGGCTGGCGGGACCTTCGCGCTTCCTGCTCGCGCTCGAGGACCTGTACAGCATC 180

QY 61 ValArgAlaAspArg 66

Db 181 GTGCGCGGTGCGACCGC 198

RESULT 8

AX395662

LOCUS AX395662 552 bp DNA linear PAT 18-MAY-2002

DEFINITION Sequence 30 from Patent WO0193897.

ACCESSION AX395662

VERSION AX395662.1 GI:21066471

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1

Sim, K. L. and Macdonald, N. J.

Angiotatin and endostatin binding proteins and methods of use

Patent: WO 0193897-A 30 13-DEC-2001;

EntreMed, Inc. (US)

Location/Qualifiers

source 1..552

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT 83 a 196 c 179 g 94 t

ORIGIN

Alignment Scores:

Pred. No.: 4,81e-17 Length: 552

Score: 26.00 Matches: 26

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 14.13% Indels: 0

DB: 6 Gaps: 0

US-09-938-391-4 (1-184) x AX395662 (1-552)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60

Db 121 GGGCTGGCGGGACCTTCGCGCTTCCTGCTCGCGCTCGAGGACCTGTACAGCATC 180

```
QY 61 ValArgAlaasparg 66
Db 181 GTGGCGCGTGGCAGCGC 198

RESULT 9
AX473835
LOCUS AX473835 552 bp DNA linear PAT 09-AUG-2002
DEFINITION Sequence 2 from Patent WO230982.
ACCESSION AX473835
VERSION AX473835.1 GI:22208005
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Sim, K.L. and Liang, H.
TITLE Angiogenesis-Inhibiting peptides and proteins and methods of use
JOURNAL Patent: WO 0230982-A 2 18-APR-2002;
EntreMed, Inc. (US)
FEATURES
source
Location/Qualifiers
1..552
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 83 a 196 c 179 g 94 t
ORIGIN
Alignment Scores:
Pred. No.: 4,81e-17 Length: 552
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 6 Gaps: 0

US-09-938-391-4 (1-184) x AX473835 (1-552)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIle 60
Db 121 GGGCTGGCGGCGACCTTCGGCGCTTCCTGCTCGCGCTGCGAGCCTGTACAGCATC 180

RESULT 10
BD081407
LOCUS BD081407 552 bp DNA linear PAT 27-AUG-2002
DEFINITION Fused protein containing angiotensin component and utilization
thereof in antitumor therapy.
ACCESSION BD081407
VERSION BD081407.1 GI:22627010
KEYWORDS JP 2001518304-A/50.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 552)
AUTHORS Bolanowski, M.A., Caparon, M.H., Casperson, G.F., Gregory, S.A.,
Klein, B.K. and McKeown, J.P.
TITLE Fused protein containing angiotensin component and utilization
thereof in antitumor therapy
JOURNAL Patent: JP 2001518304-A 50 16-OCT-2001;
GD SEARLE AND CO
COMMENT OS Homo sapiens (human)
PN JP 2001518304-A/50
PD 16-OCT-2001
PF 30-SEP-1998 JP 2000513958
PR 01-OCT-1997 US 60/060609
PI MARK A BOLANOWSKI, MAIRE H CAPARON, GERALD F CASPERSON, SUSAN A
PI GREGORY,
PI BARBARA K KLEIN, JOHN P MCKEOWN

QY 61 ValArgAlaasparg 66
Db 181 GTGGCGCGTGGCAGCGC 198

RESULT 11
AF184060
LOCUS AF184060 555 bp mRNA linear PRI 06-OCT-1999
DEFINITION Homo sapiens type XVIII collagen mRNA, complete cds.
ACCESSION AF184060
VERSION AF184060.1 GI:6013264
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 555)
AUTHORS Zhi-Yong, H., Biao, L., Wei-Jie, Z. and Xiang-Fu, W.
TITLE Cloning and expression of human endostatin gene in Escherichia coli
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 555)
AUTHORS Zhi-Yong, H., Biao, L., Wei-Jie, Z. and Xiang-Fu, W.
TITLE Direct Submission
JOURNAL Submitted (07-SEP-1999) Shanghai Institute of Biochemistry, Chinese
Academy of Sciences, 320 Yueyang Road, Shanghai 200031, P. R. China
FEATURES
source
Location/Qualifiers
1..555
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="placenta"
1..555
/note="endostatin"
/codon_start=1
/product="type XVIII collagen"
/protein_id="AAF01310.1"
/db_xref="GI:6013265"
/translation="MHSRDFQVLHLVALNSPLSGMVGIRGADFQCFQQAQAVGLA
GTFRAFLSRQLQDLYSVVRADRAAIVNLKDELLFPSEWALFSGSEGLPKGARIF
SFNGKDLVTHPTWPKQSVWHGSDPNRGLTESYCTWRTEAFSATQAYSLGLGRLIG
QSAASCHYAVILVCIENSFMFTASK"
CDS
87 a 196 c 177 g 95 t
BASE COUNT
```


Mile Road, Suite 24-26, Redford, MI 48239, USA

FEATURES

source
1..786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..786
/note="antiangiogenic agent"
/codon_start=1
/product="multi-functional protein MFP"
/protein_id="AA052249.1"
/db_xref="GI:21425751"
/translation="MSEGNSSGGSGGGVITKIKGGSGGGGGGRNPDGDAKPMWGGG
SGGGGGGYTPPKLYDGGSGGGGGSHRDQFVLHLVALNSPLSGMR
GIRGADFCQOARAVLAGTFRFLSSRLQDLYSIVRRADRAAVPIVNLKDELLFPS
WEALFSGEGPLKGPDPKGVLRHPTWPKSVWHGSDPNGRRLTESYCTWRT
EAPSATGOASSLLGRLGQSAASHAVIVLCIENSFWTASK"

CDS

BASE COUNT 124 a 239 c 280 g 143 t
ORIGIN

Alignment Scores:

Pred. No.: 6.65e-17 Length: 786
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 9 Gaps: 0

US-09-938-391-4 (1-184) x AF282883 (1-786)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60

Db 355 GGGCTGGCGGGCACCTTCGCGCTTCCTGCTCGCGCTGCAGGACCTGTACAGCATC 414

QY 61 ValArgArgAlaAspArg 66

Db 415 GTGCGCGGTGCCGACCGC 432

RESULT 15

AX370851 900 bp DNA linear PAT 01-MAR-2002

LOCUS AX370851 Sequence 12 from Patent WO0210372.

DEFINITION AX370851

ACCESSION AX370851

VERSION AX370851.1 GI:19168980

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM artificial sequences.

REFERENCE 1

AUTHORS Chapman,P.W., de Luca,G. and Falcicola,L.

TITLE Method of producing functional protein domains

JOURNAL Patent: WO 0210372-A 12 07-FEB-2002;

Applied Research Systems ARS Holding S.A. (AN)

FEATURES

source
1..900
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="mRNA for mlgSP fused to COL18A1 exons 38-41 for
pEnd-HR#1"

BASE COUNT 146 a 336 c 274 g 144 t

ORIGIN

Alignment Scores:

Pred. No.: 7.53e-17 Length: 900
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 6 Gaps: 0

US-09-938-391-4 (1-184) x AX370851 (1-900)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 457 GGGCTGGCGGGCACCTTCGCGCTTCCTGCTCGCGCTGCAGGACCTGTACAGCATC 516

QY 61 ValArgArgAlaAspArg 66

Db 517 GTGCGCGGTGCCGACCGC 534

Search completed: August 18, 2003, 02:40:06

Job time : 2930.89 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 17, 2003, 22:57:56 ; Search time 218.222 Seconds
(without alignments)
2276.105 Million cell updates/sec

Title: US-09-938-391-4
Perfect score: 184
Sequence: 1 HTHQDFQLVHLVALNSPOP.....CRHAFVLCIENSVMTSFSK 184

Scoring table:
OLIGO 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Word size: 1

Total number of hits satisfying chosen parameters: 5103490

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-Q/cn2_1/USPTO_spool_p/US09938391/runat_04082003_130738_27574/app_query.fasta_1.718
-DB=N Geneseq 19Jun03 -Qfmt=fastap -SUFFIX=p2noli.rng -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09938391 @CGN 1 1 605 @runat_04082003_130738_27574 -NCPUS=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N Geneseq 19Jun03: *
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	184	100.0	555	24	AAL46063	Canine endostatin
2	184	100.0	829	24	AAL46062	Canine pro-endosta
3	176	95.7	552	21	AAZ51309	Canine angiogenesi
4	27	14.7	555	22	AAC88291	Chicken endostatin
5	26	14.1	534	21	AAA27005	Alternate human en
6	26	14.1	537	22	AA500868	Human gene fragmen
7	26	14.1	546	21	AAA27004	Human endostatin D
8	26	14.1	549	20	AA77719	Human endostatin G
9	26	14.1	549	21	AAA29884	Human angiogenesi
10	26	14.1	549	21	AAZ51291	Human angiogenesi
11	26	14.1	549	22	AA500867	Human gene fragmen
12	26	14.1	549	25	ABA00774	Human endostatin c
13	26	14.1	551	24	ABQ81193	Human endostatin c
14	26	14.1	552	20	AA335375	SEQ ID 50 of WO991
15	26	14.1	552	21	AAC62023	Nucleotide sequenc
16	26	14.1	552	21	AA68203	Human endostatin c
17	26	14.1	552	21	AAZ50398	Human endostatin c
18	26	14.1	552	22	AAC88289	Human endostatin c
19	26	14.1	552	24	ABK50685	cDNA encoding huma
20	26	14.1	555	22	AAI66529	Human vascular end
21	26	14.1	558	24	ABA99261	Human endostatin c
22	26	14.1	563	20	AAZ08750	Human endostatin c
23	26	14.1	574	24	ABV94691	Human pancreatic c
24	26	14.1	641	22	AAH79104	Human endostatin e
25	26	14.1	641	24	AAAL5454	Specific tumour ce
26	26	14.1	641	24	AAAL44000	Reproductive recom
27	26	14.1	641	24	ABQ76079	Anticancer gene-as
28	26	14.1	650	24	ABK09978	Synthetic plasmid
29	26	14.1	816	21	ABK64013	DNA encoding a hum
30	26	14.1	900	24	ABK09977	Synthetic plasmid
31	26	14.1	968	21	AAC62025	Nucleotide sequenc
32	26	14.1	1564	24	ABQ76740	DNA encoding human
33	26	14.1	3394	18	AAT84484	Human alpha-1 coll
34	26	14.1	3394	20	AA78379	Human alpha1 (XVII
35	26	14.1	3394	24	ABN95680	Gene #2178 used to
36	26	14.1	4551	24	ABV94763	Human pancreatic c
37	26	14.1	4551	24	ABN85301	Human collagen XVI
38	26	14.1	4875	24	ABQ54955	Human ovarian anti
39	26	14.1	5408	20	AA77720	Human collagen 18
40	26	14.1	6462	21	AAA64014	Nucleotide sequenc
41	26	14.1	37664	22	AAK83781	Human immune/haema
42	26	14.1	61710	22	AAK83782	Human immune/haema
43	23	12.5	552	20	AA77715	Murine endostatin
44	23	12.5	552	21	AAZ51299	Murine angiogenesi
45	23	12.5	555	20	AA84635	Mouse endostatin c

ALIGNMENTS

RESULT 1
AAL46063
ID AAL46063 standard; cDNA; 555 BP.
XX
AC AAL46063;
XX
DT 19-JUL-2002 (first entry)
XX
DE Canine endostatin coding sequence.

Dog; pro-endostatin; endostatin; angiogenesis; cancer; gene therapy;
KW psoriasis; rheumatoid arthritis; retinopathy; macular degeneration;
KW corneal graft rejection; neovascular glaucoma; retrolental fibroplasia;
KW rubeosis; Osler-Weber Syndrome; myocardial angiogenesis;
KW plaque neovascularisation; telangiectasia; haemophilic joints;
KW angiofibroma; wound granulation; coronary collateral;
KW cerebral collateral; arteriovenous malformation;

KW ischaemic limb angiogenesis; diabetic neovascularisation; fracture;
 KW cytostatic; antiarthritic; antipsoriatic; antidiabetic; ophthalmological;
 KW synaecological; gene; ss.

XX Canis familiaris.

XX Key Location/Qualifiers
 XX CDS 1..555
 FT /*tag= a
 FT /product= "endostatin"
 FT /partial
 FT /note= "no start codon"

XX EP1191036-A2.

XX 27-MAR-2002.

XX 24-AUG-2001; 2001EP-0307224.

XX 25-AUG-2000; 2000US-227924P.

XX (PFIZ) PFIZER PROD INC.

XX Sheppard MG, Tong X;

XX WPI: 2002-354068/39.

XX P-PSDB; AAO17430.

XX An isolated nucleic acid molecule for the treatment of
 PT angiogenesis-related disorder, such as cancers or diabetic retinopathy,
 PT encodes an endostatin protein -

XX Claim 2; Fig 4; 56pp; English.

XX The present invention provides the protein and coding sequences of canine
 CC pro-endostatin and endostatin. The sequences can be used in the treatment
 CC and diagnosis of angiogenesis related disorders, including cancer,
 CC rheumatoid arthritis, psoriasis, retinopathy, macular degeneration,
 CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasia,
 CC rubeosis, Osler-Webber Syndrome, myocardial angiogenesis, plaque
 CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,
 CC wound granulation, coronary collateral, cerebral collaterals,
 CC arteriovenous malformations, ischaemic limb angiogenesis, diabetic
 CC neovascularisation, and fractures. The present sequence is the canine
 CC endostatin coding sequence.

XX SQ Sequence 555 BP; 78 A; 203 C; 191 G; 83 T; 0 other;

XX Alignment Scores:

Pred. No.: 1 55e-166 Length: 555
 Score: 184.00 Matches: 184
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

US-09-938-391-4 (1-184) x AAL46063 (1-555)

QY 1 HistThrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
 Db 1 CACACCCACCAGGACTTCACGCTGGTGTGCTGCACCTGGTGGCCCTGAACAGCCGCGCG 60
 QY 21 GlyGlyMetArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaAla 40
 Db 61 GCGCGCATGCGAGGCATCCCGGGAGCGGACTTCAGTGTCTTCAGCAGCGCGCGCGCG 120
 QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
 Db 121 GGGCTGGCGGCACCTTCGGGGCTTCCTGTGCTGGCGGTGAGACCTCTACACATC 180
 QY 61 ValArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
 Db 181 GTGCGCGCGCGCGACCGACCGGGGTGCCGTGCTCAACCTCAGGACGAGGTGCTCTTC 240

QY 81 ProSerTrpGluAlaLeuPheSerGlySerGluGlyGlnLeuValProGlyAlaArgIle 100
 Db 241 CCAGCTGGAGGCTTATTTCTCGGCTCCGAGGCCAGCTGAGCCCGGGGCGGCATC 300
 QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
 Db 301 TTCTCTTCGACGGCAGAGATGCTCTGCAGCACCCCGCTGGCCCGGGAAGAGCGTGTGG 360
 QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
 Db 361 CACGGCTCCGACCCAGCGGGCGCGCTGACGACGACGACTACTGCGAGACGTGGCGGACG 420
 QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuLeuAlaGlyArgLeuLeuGln 160
 Db 421 GAGGCCCGCGCGCCACCGGCGAGCGTCTGCTGCTGGCGGCGAGGCTGTGGAGCAG 480
 QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysIleGluAsnSerValMetThr 180
 Db 481 GAGGCCGCGAGCTGCCGCCACGCTTCGTGTGTCTGTCATCGAGAACAGCGTCATGACC 540
 QY 181 SerPheSerLys 184
 Db 541 TCCTTCTCCAAG 552

RESULT 2

ID AAL46062 standard; cDNA; 829 BP.

AC AAL46062;

DT 19-JUL-2002 (first entry)

XX Canine pro-endostatin coding sequence.

XX Dog; pro-endostatin; endostatin; angiogenesis; cancer; gene therapy;
 KW psoriasis; rheumatoid arthritis; retinopathy; macular degeneration;
 KW corneal graft rejection; neovascular glaucoma; retrolental fibroplasia;
 KW rubeosis; Osler-Webber Syndrome; myocardial angiogenesis;
 KW plaque neovascularisation; telangiectasia; haemophilic joints;
 KW angiofibroma; wound granulation; coronary collateral;
 KW cerebral collateral; arteriovenous malformation;
 KW ischaemic limb angiogenesis; diabetic neovascularisation; fracture;
 KW cytostatic; antiarthritic; antipsoriatic; antidiabetic; ophthalmological;
 KW synaecological; gene; ss.

XX Canis familiaris.

XX Key Location/Qualifiers
 FH CDS 1..693
 FT /*tag= a
 FT /product= "pro-endostatin"
 FT /partial
 FT /note= "no start codon"

XX EP1191036-A2.

XX 27-MAR-2002.

XX 24-AUG-2001; 2001EP-0307224.

XX 25-AUG-2000; 2000US-227924P.

XX (PFIZ) PFIZER PROD INC.

XX Sheppard MG, Tong X;

XX WPI: 2002-354068/39.

XX P-PSDB; AAO17429.

XX An isolated nucleic acid molecule for the treatment of
 PT angiogenesis-related disorder, such as cancers or diabetic retinopathy,
 PT encodes an endostatin protein -

XX Claim 2; Fig 2; 56pp; English.

CC The present invention provides the protein and coding sequences of canine pro-endostatin and endostatin. The sequences can be used in the treatment and diagnosis of angiogenesis related disorders, including cancer, rheumatoid arthritis, psoriasis, retinopathy, macular degeneration, corneal graft rejection, neovascular glaucoma, retrolental fibroplasia, rubeosis, Osler-Webber Syndrome, myocardial angiogenesis, plaque neovascularisation, telangiectasia, haemophilic joints, angiofibroma, wound granulation, coronary collaterals, cerebral collaterals, arteriovenous malformations, ischemic limb angiogenesis, diabetic neovascularisation, and fractures. The present sequence is the canine pro-endostatin coding sequence.

XX SQ Sequence 829 BP; 124 A; 314 C; 278 G; 113 T; 0 other;

Alignment Scores:

Pred. No.:	2.25e-166	Length:	829
Score:	184.00	Matches:	184
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-09-938-391-4 (1-184) x AAL46062 (1-829)

QY 1 HistHrHisGlnAspPheGlnLeuValLeuHisLeuValAlaLeuAsnSerProGlnPro 20
 DB 139 CACACCCACAGGACTTCCAGCTGGTGTGTCACCTGGTGGCCCTGAACAGCCGCGAGCG 198

QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnAlaArgAlaAla 40
 DB 199 GCGGCGATCCGAGGATCCCGGAGCGGACTTCCAGTGCTTCAGCAGCGCGCGCGCG 258

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerile 60
 DB 259 GGGCTGGCGCGCACCTTCCGGGCTTCTGTGTGTGGCTGTCAGGACCTCTACAGCATC 318

QY 61 ValArgArgAlaAspArgThrGlyValProValValAsnLeuArgAspGluValLeuPhe 80
 DB 319 GTGCGCGCGCGCGAGCGAGCGGCGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 378

QY 81 ProSerTrpGluAlaLeuPheSerGlySerGlyGlnLeuLysProGlyAlaArgile 100
 DB 379 CCAGCTGGAGGCTTATCTCGGCTCGAGGCGCAGCTGAGCCCGGCGCGCGCATC 438

QY 101 PheSerPheAspGlyArgAspValLeuGlnHisProAlaTrpProArgLysSerValTrp 120
 DB 439 TTCTCTTTCCAGCGCAGAGATGCTCTGTCAGCACCCTGCGCGCGGAGAGCGGTGG 498

QY 121 HisGlySerAspProSerGlyArgArgLeuThrAspSerTyrCysGluThrTrpArgThr 140
 DB 499 CACGGCTCGACCCCGAGCGGCGCGCCCTGACCGACAGCTACTGCGAGACGTGGCGGAGC 558

QY 141 GluAlaProAlaAlaThrGlyGlnAlaSerSerLeuAlaGlyArgLeuLeuGluGln 160
 DB 559 GAGGCG 618

QY 161 GluAlaAlaSerCysArgHisAlaPheValValLeuCysileGluAsnSerValMetThr 180
 DB 619 GAGGCG 678

QY 181 SerPheSerLys 184
 DB 679 TCCTTCTCCAAG 690

RESULT 3
 AAZ51309
 ID AAZ51309 standard; DNA; 552 BP.
 XX
 AC AAZ51309;
 XX

DT 06-JUN-2000 (first entry)

XX Canine angiogenesis inhibitor, endostatin DNA.

XX Canine; immunoglobulin Fc fragment; endostatin; immunofusin;
 KW angiogenesis; inhibitor; cytostatic; antirheumatoid; antiarthritic;
 KW antipsoriatic; antidiabetic; ophthalmological; immunosuppressant;
 KW vasotropic; vulnery; treatment; antiarteriosclerosis; tumour;
 KW metastasis; atherosclerosis; psoriasis; rheumatoid arthritis;
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
 KW wound granulation; keloid scar; gene therapy; ds.

XX Canis familiaris.

PH Key Location/Qualifiers
 CDS 1..552 /*tag= a
 /product= "Endostatin"
 /note= "Does not include stop codon"
 /partial

XX WO200011033-A2.

XX 02-MAR-2000.

XX 25-AUG-1999; 99WO-US19329.

XX 25-AUG-1999; 98US-0097883.

XX (LEXI-) LEXINGEN PHARM CORP.

XX Lo K, Li Y, Gillies SD;
 WIPI; 2000-237616/20.
 P-PSDB; AAY70265.

XX Novel fusion protein of angiotensin or endostatin and an immunoglobulin
 FC region, useful for treating conditions mediated by angiogenesis,
 such as rheumatoid arthritis, tumors and macular degeneration -

XX Example 8; Pages 58-59; 68pp; English.

XX The patent discloses a DNA molecule encoding a fusion protein comprising
 a signal sequence, an immunoglobulin Fc region, and an angiogenesis
 inhibitor selected from angiotensin, endostatin, a plasminogen fragment
 having angiotensin activity, a collagen XVIII fragment having endostatin
 activity, or combinations of them. The fusion protein (immunofusin) is
 used to inhibit angiogenesis and to treat diseases or conditions mediated
 by angiogenesis. Conditions that may be treated include solid tumours,
 blood born tumours, tumour metastasis, benign tumours including
 haemangiomas, acoustic neuromas, neurofibromas, trachomas and pyrogenic
 granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases
 e.g. diabetic retinopathy, retinopathy of prematurity, macular
 degeneration, corneal graft rejection, neovascular glaucoma, retrolental
 fibroplasia, rubeosis and Osler-Webber syndrome; myocardial angiogenesis,
 plaque neovascularisation, telangiectasia, haemophilic joints,
 angiofibroma, wound granulation, and excessive or abnormal stimulation of
 endothelial cells, intestinal cells, atherosclerotic, sclerodermal and
 hypertrophic scars, i.e. keloid scars. The DNA constructs may be used
 in gene therapy. The present sequence is a DNA encoding canine
 endostatin used in the construction of immunofusin containing canine
 immunoglobulin Fc fragment.

SQ Sequence 552 BP; 77 A; 204 C; 190 G; 81 T; 0 other;

Alignment Scores:

Pred. No.:	6.82e-159	Length:	552
Score:	176.00	Matches:	176
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	95.65%	Indels:	0
DB:	21	Gaps:	0

XX WPI; 2000-365617/31.
 DR P-PSDB; AAY94324.
 XX
 PT Novel endostatin capable of inhibiting endothelial cell proliferation
 PT and angiogenesis, useful for treating angiogenesis-dependent cancers
 PT and as birth control agents
 XX
 PS Claim 11; Page 40; 68pp; English.
 XX
 CC The present sequence encodes an alternate functional endostatin
 CC protein. When the human endostatin gene sequence AAA27004 is
 CC recombinantly expressed, an observable doublet of protein results, both
 CC versions of which are functional endostatin proteins. The present gene
 CC sequence encodes an endostatin variant which is the same as the protein
 CC encoded by AAA27004 minus the first four amino acids. Recombinant mouse
 CC endostatin (20 mg/kg) was administered subcutaneously to mice implanted
 CC with Lewis lung carcinomas. There was tumour mass regression
 CC non-detectable levels after 12 days of therapy due to the angiogenesis
 CC inhibitory activity of endostatin. Thus the protein is useful for
 CC treatment of angiogenesis- dependent cancers. The polynucleotide and
 CC polypeptide sequences of this endostatin are useful for treating and
 CC diagnosis of tumours, ocular angiogenic diseases, Osler-Webber syndrome,
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,
 CC haemophilic joints, angiofibroma and wound granulation, for treatment of
 CC diseases related to excessive or abnormal stimulation of endothelial
 CC cells e.g. intestinal adhesions, atherosclerosis, scleroderma. The
 CC protein may also be useful as a birth control agent by reducing or
 CC preventing uterine vascularisation. The gene for endostatin may be
 CC isolated from cells or tissue that express high levels of endostatin, eg.
 CC tumour cells, by generating cDNA from mRNA using reverse transcriptase
 CC and then amplifying the DNA sequence.
 XX
 SQ Sequence 534 BP; 77 A; 189 C; 175 G; 93 T; 0 other;

Alignment Scores:
 Pred. No.: 1.55e-15 Length: 534
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 14.13% Indels: 0
 DB: 21 Gaps: 0

US-09-938-391-4 (1-184) x AAA27005 (1-534)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
 DB 109 GGCGCTGGCGGGACCTTCCGCGCTTCTCTGCTCGCGCTGCAGGACCTGTACAGCATC 168
 QY 61 ValArgArgAlaAspArg 66
 DB 169 GTGCGCGCTGTCCGACCGC 186

RESULT 6
 AA500868
 ID AAS00868 standard; DNA; 537 BP.
 XX
 AC AAS00868;
 XX
 XX 04-JUL-2001 (first entry)
 DT
 XX Human gene fragment encoding Endostatin(TM) N-terminal mutant protein.
 DE
 XX Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;
 KW blood borne tumour; leukaemia; tumour metastasis; benign tumours;
 KW haemangioma; acoustic neuroma; neurofibroma; trachoma; rubeosis;
 KW pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
 KW retinopathy of prematurity; macular corneal graft rejection;
 KW neovascular glaucoma; retrolental fibroplasia; Osler-Webber Syndrome;
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
 KW haemophilic joint; angiofibroma; wound granulation; mutant; ds.
 XX

OS Homo sapiens.
 XX
 FH Key
 FT CDS
 FT Location/Qualifiers
 FT 1..534
 FT /*tag= a
 FT /product= "Endostatin(TM) N-terminal mutant#1"
 FT /partial
 FT /note= "Variant produced during fermentation
 FT reaction of Pichia pastoris harbouring an expression
 FT plasmid containing the present sequence"
 FT 1..540
 FT /*tag= b
 FT /product= "Endostatin(TM) N-terminal mutant#2"
 FT /partial
 FT /note= "Neither of the above CDSs have start or stop
 FT codons"
 FT CDS
 XX WO200119989-A2.
 PN 22-MAR-2001.
 PD 14-SEP-2000; 2000WO-US25166.
 PP 14-SEP-1999; 99US-0153698.
 XX (ENTR-) ENTREMED INC.
 PR Liang H, Sim KL, Chang-Murad A, Zhou X, Madsen J, Boerner RJ;
 PI Bermejo LL, Mistry PR, Shepard SR, Schrimsher JL;
 XX WPI; 2001-244802/25.
 DR P-PSDB; AAU0901, AAU0900.
 XX
 PT Producing Endostatin protein for treating angiogenesis mediated
 PT diseases such as solid tumours, comprises recombinantly producing the
 PT protein using an expression system, and recovering and purifying the
 PT protein
 XX
 PS Claim 6; Page 32; 67pp; English.
 XX
 CC The sequence encodes Human Endostatin(TM) N-terminal deletion
 CC mutant lacking the N-terminal 4 amino acids. The new method of the
 CC invention is useful for producing, recovering and purifying Endostatin
 CC (TM) from biological sources, such as biological fluids, tissues, cells,
 CC culture media, and fermentation media. Endostatin(TM) is useful for
 CC treating angiogenesis mediated diseases such as solid tumours, blood
 CC borne tumours, leukaemias, tumour metastases, benign tumours, e.g.
 CC haemangioma, acoustic neuromas, neurofibromas, trachomas, and pyogenic
 CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases,
 CC e.g., diabetic retinopathy, retinopathy of prematurity, macular
 CC degeneration, corneal graft rejection, neovascular glaucoma, colon
 CC cancer, retrolental fibroplasia, rubeosis, Osler-Webber Syndrome,
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,
 CC haemophilic joints, angiofibroma, and wound granulation. Endostatin(TM)
 CC is also useful for treating disease of excessive or abnormal stimulation
 CC of endothelial cells such as intestinal adhesions, atherosclerosis,
 CC scleroderma and hypertrophic scars. Higher yields of more purified, and
 CC biologically active Endostatin(TM) are obtained by the new method.
 CC Endostatin(TM) can be stored in buffers for extended periods of time, and
 CC also subjected to lyophilisation, while preserving biological activity.
 CC Centrifugation of broth from fermentation steps in production is avoided,
 CC preventing unwanted potential cellular lysis and contamination with
 CC additional proteins, pigments, enzymes and other cellular chemicals and
 CC debris.
 XX
 SQ Sequence 537 BP; 79 A; 189 C; 176 G; 93 T; 0 other;

Alignment Scores:
 Pred. No.: 1.56e-15 Length: 537
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 14.13% Indels: 0


```

XX SQ Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;
Alignment Scores:
Pred. No.: 1-59e-15 Length: 549
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 20 Gaps: 0

US-09-938-391-4 (1-184) x AAX77719 (1-549)
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 121 GGGCTGGCGGCACCTTCGCGCCTTCCTCGCGCTGCAGGACCTGTACAGCATC 180
QY 61 ValArgArgAlaAspArg 66
Db 181 GTGCGCGCTGCCGACCGC 198

RESULT 9
AAA29884
ID AAA29884 standard; cDNA; 549 BP.
XX
AC AAA29884;
DT 22-AUG-2000 (first entry)
XX
DE Human angiogenesis inhibiting factor 1 encoding cDNA.
KW Human; angiogenesis inhibiting factor 1; IAF-1; tumour; antibody;
KW abnormal vessel disease; ss.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
CDS 1..549
FT /*tag= a
FT /product= "Endostatin"
FT /note= "Does not include stop codon"
FT /partial
XX
PN WO200011033-A2.
XX
PD 02-MAR-2000.
XX
PF 25-AUG-1999; 99WO-US19329.
XX
PR 25-AUG-1998; 98US-0097883.
XX
PA (LEXI-) LEXINGEN PHARM CORP.
XX
PI Lo K, Li Y, Gillies SD;
XX
WPI; 2000-237616/20.
DR P-PSDB; AAY70252.
XX
PT Novel fusion protein of angiostatin or endostatin and an immunoglobulin
XX FC region, useful for treating conditions mediated by angiogenesis,
XX such as rheumatoid arthritis, tumors and macular degeneration -
XX
PS Claim 12; Pages 40-41; 68pp; English.
XX
CC The patent discloses a DNA molecule encoding a fusion protein comprising
XX a signal sequence, an immunoglobulin FC region, and an angiogenesis
XX inhibitor selected from angiostatin, endostatin, a plasminogen fragment
XX having angiostatin activity, a collagen XVIII fragment having endostatin
XX activity, or combinations of them. The fusion protein (immunofusin) is
XX used to inhibit angiogenesis and to treat diseases or conditions mediated
XX by angiogenesis. Conditions that may be treated include solid tumours,
XX blood born tumours, tumour metastasis, benign tumours including
XX haemangiomas, acoustic neuromas, neurofibromas, trachomas and pyrogenic
XX granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases
XX e.g. diabetic retinopathy, retinopathy of prematurity, macular
XX degeneration, corneal graft rejection, neovascular glaucoma, retrolental

```

CC fibroplasia, rubeosis and Osler-webber syndrome; myocardial angiogenesis,
CC plaque neovascularisation, telangiectasia, haemophilic joints,
CC angiofibroma, wound granulation, and excessive or abnormal stimulation of
CC endothelial cells, intestinal cells, atherosclerosis, scleroderma and
CC hypertrophic scars, i.e. keloid scars. The DNA constructs may be used
CC in gene therapy. The present sequence is a cDNA encoding human
CC endostatin used in the construction of immunofusin containing human
CC immunoglobulin gamma (IgG) Fc fragment.
CC Note: This sequence is stated in Claim 12 as being amino acid sequence
CC of plasminogen fragment, however, the rest of the specification
CC refers to this sequence as being nucleotide sequence of human endostatin.
XX
SQ Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;

Alignment Scores:
Pred. No.: 1.59e-15 Length: 549
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 21 Gaps: 0

US-09-938-391-4 (1-184) x AAZ51291 (1-549)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
DB 121 GGGCTGGCGGGACCTTCGGCGCTTCCTGCTCGCGCTGAGGACCTGTACAGCATC 180
QY 61 ValArgAlaAspArg 66
DB 181 GTGCGCGGTGGCAGCGC 198

RESULT 11
AAS00867
ID AAS00867 standard; DNA; 549 BP.

XX AAS00867;
XX
XX
DT 04-JUL-2001 (first entry)
XX
DE Human gene fragment encoding Endostatin(TM) protein.
XX Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;
KW blood borne tumour; leukaemia; tumour metastasis; benign tumour;
KW haemangioma; acoustic neuroma; neurofibroma; trachoma; rubeosis;
KW pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;
KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
KW retinopathy of prematurity; macular corneal graft rejection;
KW neovascular glaucoma; retrolental fibroplasia; Osler-Webber Syndrome;
KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
KW haemophilic joint; angiofibroma; wound granulation; ds.

OS Homo sapiens.

XX
XX Key Location/Qualifiers
FT 1..540
FT CDS
FT /tag= a
FT /product= "Endostatin(TM) C-terminus minus 3"
FT /partial
FT /note= "Variant produced during fermentation
FT reaction of Pichia pastoris harbouring an expression
FT plasmid containing the present sequence"
FT 1..543
FT CDS
FT /tag= b
FT /product= "Endostatin(TM) C-terminus minus 2"
FT /partial
FT /note= "Variant produced during fermentation
FT reaction of Pichia pastoris harbouring an expression
FT plasmid containing the present sequence"
FT 1..546
FT CDS
FT /tag= c
FT /product= "Endostatin(TM) C-terminus minus 1"
FT /partial

FT
FT
FT
FT CDS
FT /note= "Variant produced during fermentation
FT reaction of Pichia pastoris harbouring an expression
FT plasmid containing the present sequence"
FT 1..549
FT /tag= d
FT /product= "Endostatin(TM)"
FT /partial
FT /note= "None of the above CDSs have start or stop codons"
XX
XX WO200119989-A2.
XX
XX 22-MAR-2001.
XX
XX 14-SEP-2000; 2000WO-US25166.
XX
XX 14-SEP-1999; 99US-0153698.
XX (ENTR-) ENTREMED INC.
XX Liang H, Sim KL, Chang-Murad A, Zhou X, Madsen J, Boerner RJ;
XX Bermejo LL, Mistry FR, Shepard SR, Schrimsher JL;
XX
XX WPI; 2001-244802/25.
XX P-PSDB; AAU00896, AAU00897, AAU00898, AAU00899.

XX Producing Endostatin protein for treating angiogenesis mediated
XX diseases such as solid tumours, comprises recombinantly producing the
XX protein using an expression system, and recovering and purifying the
XX protein
XX
XX Claim 6; Page 29; 67pp; English.

XX The sequence encodes Human Endostatin(TM). The new method of the
XX invention is useful for producing, recovering and purifying Endostatin
XX (TM) from biological sources, such as biological fluids, tissues, cells,
XX culture media, and fermentation media. Endostatin(TM) is useful for
XX treating angiogenesis mediated diseases such as solid tumours, blood
XX borne tumours, leukaemias, tumour metastases, benign tumours, e.g.
XX haemangioma, acoustic neuromas, neurofibromas, trachomas, and pyogenic
XX granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases,
XX e.g., diabetic retinopathy, retinopathy of prematurity, macular
XX degeneration, corneal graft rejection, neovascular glaucoma, colon
XX cancer, retrolental fibroplasia, rubeosis, Osler-Webber Syndrome,
XX myocardial angiogenesis, plaque neovascularisation, telangiectasia,
XX haemophilic joints, angiofibroma, and wound granulation. Endostatin(TM)
XX is also useful for treating disease of excessive or abnormal stimulation
XX of endothelial cells such as intestinal adhesions, atherosclerosis, and
XX scleroderma and hypertrophic scars. Higher yields of more purified, and
XX biologically active Endostatin(TM) are obtained by the new method.
XX Endostatin(TM) can be stored in buffers for extended periods of time, and
XX also subjected to lyophilisation, while preserving biological activity.
XX Centrifugation of broth from fermentation steps in production is avoided,
XX preventing unwanted potential cellular lysis and contamination with and
XX additional proteins, pigments, enzymes and other cellular chemicals and
XX debris.

XX Sequence 549 BP; 82 A; 196 C; 178 G; 93 T; 0 other;

Alignment Scores:

Pred. No.: 1.59e-15 Length: 549
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 22 Gaps: 0

US-09-938-391-4 (1-184) x AAS00867 (1-549)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
DB 121 GGGCTGGCGGGACCTTCGGCGCTTCCTGCTCGCGCTGAGGACCTGTACAGCATC 180
QY 61 ValArgAlaAspArg 66

CC sarcoidosis, serpinginous or geographic choroiditis, subretinal
CC fluid drainage, tiled disc syndrome, Toxoplasma retinochoroiditis,
CC tuberculosis, Vogt-Koyanagi-Harada syndrome, diabetic retinopathy,
CC non-diabetic retinopathy, brain vein occlusion, central retinal
CC vein occlusion, retinopathy in premature infants, rubeosis iridis,
CC neovascular glaucoma, perifoveal telangiectasis, sickle cell
CC retinopathy, Eale's disease, retinal vasculitis, Von Hippel
CC Lindau disease, radiation retinopathy, retinal cryoinjury,
CC retinitis pigmentosa, retinochoroidal coloboma, corneal
CC neovascularisation due to herpes simplex keratitis, corneal ulcers,
CC keratoplasty, pterygia and trauma (all claimed).
XX
XX Sequence 551 BP; 83 A; 195 C; 179 G; 94 T; 0 other;
XX

Alignment Scores:
Pred. No.: 1.6e-15 Length: 551
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 24 Gaps: 0

US-09-938-391-4 (1-184) x ABQ81193 (1-551)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIle 60
DB 120 GGGCTGGCGGGACCTTCGGCGCTTCCTGTCTCGCGCTGCAGGACCTGTACAGCATC 179

QY 61 ValArgArgAlaAspArg 66
DB 180 GTGCGCGCTGCCGACCGC 197

RESULT 14
AAX35375
ID AAX35375 standard; DNA; 552 BP.
XX
AC AAX35375;
DT 16-JUL-1999 (first entry)
XX
DE SEQ ID 50 of WO9916889.
XX
KW Angiostatin; endostatin; interferon; thrombospondin;
KW interferon-inducible protein; platelet factor 4; anti-angiogenic;
KW anti-tumor; multifunctional protein; angiogenic-mediated disease;
KW cancer; diabetic retinopathy; macular degeneration; arthritis;
KW tumor cell production; ss.
XX
OS Homo sapiens.
XX
PN WO9916889-A1.
XX
PD 08-APR-1999.
XX
PF 30-SEP-1998; 98WO-US20464.
XX
PR 01-OCT-1997; 97US-0060609.
XX
XX (SEAR) SEARLE & CO G D.
XX
XX Bolanowski MA, Caparon MH, Casperson GF, Gregory SA;
PI Klein BK, McKearn JP;
XX
XX WPI; 1999-255098/21.
XX
XX New multifunctional proteins useful for treating angiogenic-mediated
XX diseases
XX
XX Disclosure; Page 85; 121pp; English.
XX
XX The specification describes multifunctional proteins which comprise
XX combinations of angiostatin, endostatin, interferon, thrombospondin,
XX interferon-inducible protein and platelet factor 4, and have

CC anti-angiogenic and/or anti-tumor activity. The multifunctional protein
CC may exhibit useful properties such as having similar or greater
CC biological activity when compared to a single factor or by having
CC improved half-life or decreased adverse side effects, or a combination
CC of these properties. The proteins can be used for treating an
CC angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular
CC degeneration, or arthritis. They can also be used for inhibiting the
CC production of tumor cells (characteristic of lung, breast, ovarian,
CC prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma,
CC hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor
CC growth. The present sequence is used in the course of the invention.
XX
XX Sequence 552 BP; 83 A; 196 C; 179 G; 94 T; 0 other;
XX

Alignment Scores:
Pred. No.: 1.6e-15 Length: 552
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 20 Gaps: 0

US-09-938-391-4 (1-184) x AAX35375 (1-552)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIle 60
DB 121 GGGCTGGCGGGACCTTCGGCGCTTCCTGTCTCGCGCTGCAGGACCTGTACAGCATC 180

QY 61 ValArgArgAlaAspArg 66
DB 181 GTGCGCGCTGCCGACCGC 198

RESULT 15
AAC62023
ID AAC62023 standard; DNA; 552 BP.
XX
AC AAC62023;
DT 06-MAR-2001 (first entry)
XX
DE Nucleotide sequence of human endostatin encoded by plasmid pVALch#15.
XX
KW Streptomyces sp. strain C5; SnpA; S. venezuelae; alpha-amylase;
KW endostatin; cancer; tumour growth; angiogenesis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..552
FT /tag= a
FT /product= "endostatin"
XX
PN WO2000060945-A1.
XX
PD 19-OCT-2000.
XX
PF 12-APR-2000; 2000WO-US09747.
XX
PR 13-APR-1999; 99US-0129084.
XX
XX (MERI) MERCK & CO INC.
XX
XX Desanti CL, Strohl WR;
XX
XX WPI; 2000-686970/67.
XX
XX P-PSDB; AAB30493.
XX
XX Preparation of soluble recombinant endostatin involves transforming
XX Streptomyces host with expression vector comprising nucleotide
XX sequence encoding endostatin operably linked to linker and leader
XX peptide
XX
XX Example 1; Fig 6; 57pp; English.
PS

XX The present sequence encodes human endostatin. The protein is expressed
CC in Streptomyces. Leader sequences of Streptomyces sp. strain C5
CC SnpA and S. venezuelae alpha-amylase proteins are linked to the
CC N-terminal of endostatin. This ensures that endostatin protein is
CC produced as a secreted, soluble protein which needs no refolding, is
CC stable in the fermentation broth and is produced in large quantities.
CC The method is used for preparing soluble recombinant human, murine or
CC primate endostatin, which is useful in the treatment of cancer,
CC inhibition of tumour growth, inhibition of angiogenesis, isolation of
CC receptors for endostatin and for identification of anti-angiogenic
CC compounds in assays. The endostatin protein is produced as a secreted,
CC soluble protein which needs no refolding, is stable in the fermentation
CC broth and is produced in large quantities. Streptomyces are amenable
CC for cultivation in large fermentations allowing for large quantities of
CC soluble endostatin to be produced.
XX
SQ Sequence 552 BP; 83 A; 196 C; 179 G; 94 T; 0 other;

Alignment Scores:
Pred. No.: 1.6e-15 Length: 552
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 21 Gaps: 0

US-09-938-391-4 (1-184) x AAC62023 (1-552)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 121 GGCGTGGCGGGCACCTTCGCGGCTTCCTGTCCTCGCGCTGCAGGACCTGTACAGCATC 180

QY 61 ValArgArgAlaAspArg 66
Db 181 GTGCGCCGTGCGGACCGC 198

Search completed: August 18, 2003, 00:50:04
Job time : 220.222 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 18, 2003, 00:40:31 ; Search time 57.3333 Seconds
(without alignments)
1416.532 Million cell updates/sec

Title: US-09-938-391-4
Perfect score: 184
Sequence: 1 HTHQDFQLVHLVALNSQP.....CRHAFVLCIENSVMTSFSK 184

Scoring table:
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1135299

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xl
-Q=/cgn2_1/usfto.spool_p/US09938391/runat_04082003_130739_27621/app.query.fasta_1.718
-DB=Issued Patents NA -OPT=fastap -SUFFIX=p2noli.rn1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938391 @CGN_1_1_133 @runat_04082003_130739_27621 -NCPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/prodata/2/ina/5A COMB.seq:
2: /cgn2_6/prodata/2/ina/5B COMB.seq:
3: /cgn2_6/prodata/2/ina/6A COMB.seq:
4: /cgn2_6/prodata/2/ina/6B COMB.seq:
5: /cgn2_6/prodata/2/ina/PTUS COMB.seq:
6: /cgn2_6/prodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	14.1	534	US-09-315-689-6	Sequence 6, Appli
2	26	14.1	546	US-09-315-689-4	Sequence 4, Appli
3	26	14.1	552	US-09-206-059-30	Sequence 30, Appli
4	26	14.1	3394	US-08-159-784-4	Sequence 4, Appli
5	23	12.5	558	US-09-449-293-3	Sequence 3, Appli
6	23	12.5	558	US-09-775-325-3	Sequence 3, Appli
7	23	12.5	565	US-08-985-526-37	Sequence 37, Appli
8	23	12.5	573	US-09-561-500-12	Sequence 12, Appli
9	23	12.5	573	US-09-561-108-12	Sequence 12, Appli
10	23	12.5	573	US-09-561-526-12	Sequence 12, Appli
11	23	12.5	573	US-09-561-499-12	Sequence 12, Appli
12	23	12.5	4031	US-08-159-784-1	Sequence 1, Appli

c 13	9	4.9	1823	4	US-09-620-312D-1012	Sequence 1012, Ap
c 14	8	4.3	508	1	US-08-318-193-69	Sequence 69, Appl
c 15	8	4.3	800	4	US-09-222-938A-68	Sequence 68, Appl
c 16	8	4.3	1001	4	US-09-641-638-113	Sequence 113, App
c 17	8	4.3	1001	4	US-09-641-638-114	Sequence 114, App
c 18	8	4.3	1305	4	US-09-107-532A-1972	Sequence 1972, Ap
c 19	8	4.3	4748	3	US-09-331-581-1	Sequence 1, Appli
c 20	8	4.3	5798	4	US-09-595-386-1	Sequence 3, Appli
c 21	8	4.3	7286	3	US-09-331-581-3	Sequence 14, Appl
c 22	8	4.3	7938	3	US-09-331-581-14	Sequence 207, App
c 23	8	4.3	29604	3	US-08-781-891-207	Sequence 207, App
c 24	8	4.3	29604	4	US-09-618-168-207	Sequence 207, App
c 25	8	4.3	111282	4	US-09-754-250-3	Sequence 3, Appli
c 26	7	3.8	32	4	US-09-522-217-69	Sequence 69, Appl
c 27	7	3.8	33	1	US-08-368-803-27	Sequence 27, Appl
c 28	7	3.8	33	2	US-08-578-096A-28	Sequence 28, Appl
c 29	7	3.8	33	3	US-09-240-426-28	Sequence 28, Appl
c 30	7	3.8	40	1	US-08-162-597-5	Sequence 5, Appli
c 31	7	3.8	45	3	US-09-232-478-26	Sequence 26, Appl
c 32	7	3.8	45	4	US-09-785-055-26	Sequence 26, Appl
c 33	7	3.8	70	1	US-08-434-001-98	Sequence 98, Appl
c 34	7	3.8	70	1	US-08-433-585-98	Sequence 98, Appl
c 35	7	3.8	70	1	US-08-434-425-98	Sequence 98, Appl
c 36	7	3.8	70	2	US-08-437-667-98	Sequence 98, Appl
c 37	7	3.8	70	3	US-08-906-955-98	Sequence 98, Appl
c 38	7	3.8	70	3	US-08-945-909-98	Sequence 98, Appl
c 39	7	3.8	70	4	US-09-396-002A-98	Sequence 98, Appl
c 40	7	3.8	70	5	PCT-US96-06050-98	Sequence 98, Appl
c 41	7	3.8	71	1	US-08-434-001-97	Sequence 97, Appl
c 42	7	3.8	71	1	US-08-433-585-97	Sequence 97, Appl
c 43	7	3.8	71	1	US-08-434-425-97	Sequence 97, Appl
c 44	7	3.8	71	2	US-08-437-667-97	Sequence 97, Appl
c 45	7	3.8	71	3	US-08-906-955-97	Sequence 97, Appl

ALIGNMENTS

RESULT 1
US-09-315-689-6
; Sequence 6, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-315-689-6

Alignment Scores:
Pred. No.: 1.08e-16 Length: 534
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 4 Gaps: 0

US-09-938-391-4 (1-184) x US-09-315-689-6 (1-534)

Qy	41	GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrsIle 60
Db	109	GGGCTGGCGGGACCTTCGGCGCTCTCTGCTCGCGCTGCAGGACCTGTACAGCA 168
Qy	61	ValArgAlaAspArg 66
Db	169	GTGGCGCGTGGCGGACCG 186

```

RESULT 2
US-09-315-689-4
; Sequence 4, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Folkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-315-689-4
Alignment Scores:
Pred. No.: 1,1e-16 Length: 546
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
Gaps: 4
DB:

US-09-938-391-4 (1-184) x US-09-315-689-4 (1-546)
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 121 GGGCTGGGGGACCTTCGGCGCTTCCTGCTCGCGCTGCAGGACCTGTACAGCATC 180
QY 61 ValArgAlaAspArg 66
Db 181 GTGCGCGCTGCGCAGCGC 198

RESULT 3
US-09-206-059-30
; Sequence 30, Application US/09206059
; Patent No. 6201104
; GENERAL INFORMATION:
; APPLICANT: MacDonald, Nicholas
; APPLICANT: Sim, Kim Lee
; TITLE OF INVENTION: Angiogenesis-Inhibiting Protein Binding Peptides and
; FILE REFERENCE: 05213-0370
; CURRENT APPLICATION NUMBER: US/09/206,059
; CURRENT FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-206-059-30
Alignment Scores:
Pred. No.: 1,11e-16 Length: 552
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
Gaps: 3
DB:

US-09-938-391-4 (1-184) x US-09-206-059-30 (1-552)
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 121 GGGCTGGGGGACCTTCGGCGCTTCCTGCTCGCGCTGCAGGACCTGTACAGCATC 180
QY 61 ValArgAlaAspArg 66

Db 181 GTGCGCGCTGCGCAGCGC 198

RESULT 4
US-08-159-784-4
; Sequence 4, Application US/08159784
; Patent No. 5643783
; GENERAL INFORMATION:
; APPLICANT: Bjorn R. Olsen
; TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,784
; FILING DATE: December 1, 1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: John F. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00246/170001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3394
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-159-784-4
Alignment Scores:
Pred. No.: 6,02e-16 Length: 3394
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
Gaps: 1
DB:

US-09-938-391-4 (1-184) x US-08-159-784-4 (1-3394)
QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 1624 GGGCTGGGGGACCTTCGGCGCTTCCTGCTCGCGCTGCAGGACCTGTACAGCATC 1683
QY 61 ValArgAlaAspArg 66
Db 1684 GTGCGCGCTGCGCAGCGC 1701

RESULT 5
US-09-449-293-3
; Sequence 3, Application US/09449293
; Patent No. 6267954
; GENERAL INFORMATION:
; APPLICANT: Abitbol, Marc
; APPLICANT: Uteza, Yves
; APPLICANT: Menasche, Maurice
; APPLICANT: Bossard, Carine
```


APPLICANT: Van Den Bergh, Loic
APPLICANT: Bonnel, Sebastien
APPLICANT: Prats, Herve
APPLICANT: Honiger, Jiri
APPLICANT: Neuner-Jehle, Martin
TITLE OF INVENTION: INTRAOCULAR TRANSPLANTATION OF ENCAPSULATED CELLS
FILE REFERENCE: 8076.202US01
CURRENT APPLICATION NUMBER: US/09/449,293
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 558
TYPE: DNA
ORGANISM: Rattus rattus
US-09-449-293-3

Alignment Scores:
Pred. No.: 9.69e-14 Length: 558
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.50% Indels: 0
DB: 3 Gaps: 0

US-09-938-391-4 (1-184) x US-09-449-293-3 (1-558)

QY 44 GlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIleValArgArg 63
Db 133 GGACCTTCGGGGCTTCTCTAGGCTGCAGGATCTCTATAGCATCGTGCGCGT 192

QY 64 AlaAspArg 66
Db 193 GCTGACCGG 201

RESULT 6
US-09-775-325-3
Sequence 3, Application US/09775325
Patent No. 6500449
GENERAL INFORMATION:
APPLICANT: Abitbol, Marc
APPLICANT: Uteza, Yves
APPLICANT: Menasche, Maurice
APPLICANT: Bossard, Carine
APPLICANT: Van Den Bergh, Loic
APPLICANT: Bonnel, Sebastien
APPLICANT: Prats, Herve
APPLICANT: Honiger, Jiri
APPLICANT: Neuner-Jehle, Martin
TITLE OF INVENTION: INTRAOCULAR TRANSPLANTATION OF ENCAPSULATED CELLS
FILE REFERENCE: 8076.202US01
CURRENT APPLICATION NUMBER: US/09/775,325
CURRENT FILING DATE: 2001-02-01
PRIOR APPLICATION NUMBER: US 09/449,293
PRIOR FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 558
TYPE: DNA
ORGANISM: Rattus rattus
US-09-775-325-3

Alignment Scores:
Pred. No.: 9.69e-14 Length: 558
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.50% Indels: 0
DB: 4 Gaps: 0

US-09-938-391-4 (1-184) x US-09-775-325-3 (1-558)

QY 44 GlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIleValArgArg 63
Db 133 GGACCTTCGGGGCTTCTCTAGGCTGCAGGATCTCTATAGCATCGTGCGCGT 192

QY 64 AlaAspArg 66
Db 193 GCTGACCGG 201

RESULT 7
US-08-985-526-37
Sequence 37, Application US/08985526
Patent No. 6080728
GENERAL INFORMATION:
APPLICANT: Mixson, James A
TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA
TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
TITLE OF INVENTION: THERAPY
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Connolly, Bove, Lodge, & Hutz
STREET: 1220 Market Street, P.O. Box 2207
CITY: Wilmington
STATE: Delaware
COUNTRY: U.S.A.
ZIP: 19899
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,526
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,845
FILING DATE: 16-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: McMorro Jr., Robert G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (302) 658-9141
TELEFAX: (302) 658-5613
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 565 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-985-526-37

Alignment Scores:
Pred. No.: 9.8e-14 Length: 565
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.50% Indels: 0
DB: 3 Gaps: 0

US-09-938-391-4 (1-184) x US-08-985-526-37 (1-565)

QY 44 GlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIleValArgArg 63
Db 140 GGACCTTCGGGGCTTCTCTAGGCTGCAGGATCTCTATAGCATCGTGCGCGT 199

QY 64 AlaAspArg 66
Db 200 GCTGACCGG 208

RESULT 8
US-09-561-500-12
Sequence 12, Application US/09561500
Patent No. 6342219
GENERAL INFORMATION:

APPLICANT: Philip E. Thorpe
APPLICANT: Rolf A. Brekken
TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
FILE REFERENCE: 4001.002500
CURRENT APPLICATION NUMBER: US/09/561,500
CURRENT FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/131,432
PRIOR FILING DATE: 1999-04-28
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 573
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
OTHER INFORMATION: OLIGONUCLEOTIDE
NAME/KEY: CDS
LOCATION: (1)..(573)
US-09-561-500-12

Alignment Scores:
Pred. No.: 9.93e-14 Length: 573
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.50% Indels: 0
Gaps: 4
DB:

US-09-938-391-4 (1-184) x US-09-561-500-12 (1-573)

QY 44 GlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTySerIleValArg 63
DB 151 GGCACCTTCGGGCTTCTCTAGGCTGCAGGATCTCTATAGCATCGTGCCCGT 210

QY 64 AlaAspArg 66
DB 211 GCTGACCGG 219

RESULT 9
US-09-561-108-12
Sequence 12, Application US/09561108
Patent No. 6342221
GENERAL INFORMATION:
APPLICANT: Philip E. Thorpe
APPLICANT: Rolf A. Brekken
TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
FILE REFERENCE: 4001.002584
CURRENT APPLICATION NUMBER: US/09/561,108
CURRENT FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/131,432
PRIOR FILING DATE: 1999-04-28
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 573
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
OTHER INFORMATION: OLIGONUCLEOTIDE
NAME/KEY: CDS
LOCATION: (1)..(573)
US-09-561-108-12

Alignment Scores:
Pred. No.: 9.93e-14 Length: 573
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.50% Indels: 0
Gaps: 4
DB:

US-09-938-391-4 (1-184) x US-09-561-108-12 (1-573)

QY 44 GlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTySerIleValArg 63
DB 151 GGCACCTTCGGGCTTCTCTAGGCTGCAGGATCTCTATAGCATCGTGCCCGT 210

QY 64 AlaAspArg 66
DB 211 GCTGACCGG 219

RESULT 10
US-09-561-526-12
Sequence 12, Application US/09561526
Patent No. 6416758
GENERAL INFORMATION:
APPLICANT: Philip E. Thorpe
APPLICANT: Rolf A. Brekken
TITLE OF INVENTION: ANTIBODY CONJUGATE KITS FOR SELECTIVELY INHIBITING VEGF
FILE REFERENCE: 4001.002586
CURRENT APPLICATION NUMBER: US/09/561,526
CURRENT FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/131,432
PRIOR FILING DATE: 1999-04-28
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 573
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
OTHER INFORMATION: OLIGONUCLEOTIDE
NAME/KEY: CDS
LOCATION: (1)..(573)
US-09-561-526-12

Alignment Scores:
Pred. No.: 9.93e-14 Length: 573
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.50% Indels: 0
Gaps: 4
DB:

US-09-938-391-4 (1-184) x US-09-561-526-12 (1-573)

QY 44 GlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTySerIleValArg 63
DB 151 GGCACCTTCGGGCTTCTCTAGGCTGCAGGATCTCTATAGCATCGTGCCCGT 210

QY 64 AlaAspArg 66
DB 211 GCTGACCGG 219

RESULT 11
US-09-561-499-12
Sequence 12, Application US/09561499
Patent No. 6524583
GENERAL INFORMATION:
APPLICANT: Philip E. Thorpe
APPLICANT: Rolf A. Brekken
TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF
FILE REFERENCE: 4001.002582
CURRENT APPLICATION NUMBER: US/09/561,499
CURRENT FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/131,432
PRIOR FILING DATE: 1999-04-28
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 573
TYPE: DNA
ORGANISM: Artificial Sequence

FEATURE: OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
OTHER INFORMATION: OLIGONUCLEOTIDE
NAME/KEY: CDS
LOCATION: (1) ..(573)
US-09-561-499-12

Alignment Scores:
Pred. No.: 9.93e-14 Length: 573
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.50% Indels: 0
DB: 4 Gaps: 0

US-09-938-391-4 (1-184) x US-09-561-499-12 (1-573)

QY 44 GlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTy-SerIleValArgArg 63
DB 151 GGCACCTTCGGGCTTTCCTGCTAGGCTGCAGGATCTATAGCATCGTGGCGGT 210
QY 64 AlaAspArg 66
DB 211 GCTGACCGG 219

RESULT 12
US-08-159-784-1
Sequence 1, Application US/08159784
Patent No. 5643783
GENERAL INFORMATION:
APPLICANT: Bjorn R. Olsen
TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/159,784
FILING DATE: December 1, 1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: John F. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00246/170001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4031
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-159-784-1

Alignment Scores:
Pred. No.: 6.09e-13 Length: 4031
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 12.50% Indels: 0
DB: 1 Gaps: 0

US-09-938-391-4 (1-184) x US-08-159-784-1 (1-4031)

QY 44 GlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTy-SerIleValArgArg 63
DB 3442 GGCACCTTCGGGCTTTCCTGCTAGGCTGCAGGATCTATAGCATCGTGGCGGT 3501
QY 64 AlaAspArg 66
DB 3502 GCTGACCGG 3510

RESULT 13
US-09-620-312D-1012/c
Sequence 1012, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 1012
LENGTH: 1823
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (164)..(1786)
US-09-620-312D-1012

Alignment Scores:
Pred. No.: 14.5 Length: 1823
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.89% Indels: 0
DB: 4 Gaps: 0

US-09-938-391-4 (1-184) x US-09-620-312D-1012 (1-1823)

QY 40 AlaGlyLeuAlaGlyThrPheArgAla 48
DB 1113 GCTGGTTGGCAGGCACCTTCAGGGCC 1087

RESULT 14
US-08-318-193-69/c
Sequence 69, Application US/08318193
Patent No. 5641663

GENERAL INFORMATION:
; APPLICANT: GARVIN, Robert T.
; APPLICANT: MALEK, Lawrence T.
; TITLE OF INVENTION: AN EXPRESSION SYSTEM FOR THE SECRETION
; TITLE OF INVENTION: OF BIOACTIVE HUMAN GRANULOCYTE MACROPHAGE COLONY
; TITLE OF INVENTION: STIMULATING FACTOR (GM-CSF) AND OTHER HETEROLOGOUS
; TITLE OF INVENTION: PROTEINS FROM STREPTOMYCES
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,193
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,314
; FILING DATE:
; APPLICATION NUMBER: US 07/224,568
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 18740/116 CACO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 508 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Other nucleic acid;
; DESCRIPTION: Synthetic DNA oligonucleotide
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5..502
; US-08-318-193-69

Alignment Scores:
Pred. No.: 42 82 Length: 508
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.35% Indels: 0
DB: 1 Gaps: 0

US-09-938-391-4 (1-184) x US-08-318-193-69 (1-508)

QY 37 AlaAgaAlaAlaGlyLeuAlaGly 44
|||||
DB 496 GCCCGTGGCGAGGCGCTCGCGGT 473

RESULT 15
US-09-222-938A-68
; Sequence 68, Application US/09222938A
; Patent No. 6437108
; GENERAL INFORMATION:
; APPLICANT: Youngman, Philip
; APPLICANT: Fritz, Christian
; APPLICANT: Murphy, Christopher
; APPLICANT: Guzman, Luz-Maria
; TITLE OF INVENTION: ESSENTIAL BACTERIAL GENES AND THEIR USE

FILE REFERENCE: 07334/060001
; CURRENT APPLICATION NUMBER: US/09/222,938A
; CURRENT FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 68
; LENGTH: 800
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (114)...(743)
; US-09-222-938A-68
Alignment Scores:
Pred. No.: 64.3 Length: 800
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.35% Indels: 0
DB: 4 Gaps: 0
US-09-938-391-4 (1-184) x US-09-222-938A-68 (1-800)
QY 156 ArgLeuLeuGluGlnGluAlaAla 163
|||||
DB 760 AGGCTATTAGACCAAGACGACGA 783
Search completed: August 18, 2003, 03:44:02
Job time : 61.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 18, 2003, 02:40:12 ; Search time 725.333 Seconds
(without alignments)
567.713 Million cell updates/sec

Title: US-09-938-391-4
Perfect score: 184
Sequence: 1 HTHQDFQLVHLVALNSQP.....CRHAFVVLCIENSVMTSFSK 184

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1504479 seqs, 1118970152 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3005078

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -DEV-xlp
-Q/cgn2_1/USPTO_spool_p/US0938391/runat_04082003_130741_27700/app_query.fasta_1.718
-DB=Published Applications NA -QWTF=fastcap -SUFFIX=p2noli.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cd -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ext -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USRP=US0938391@cgn 1.129 @runat_04082003_130741_27700
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=6 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB	ID	Description
1	176	95.7	552	12	US-10-292-418-34	Sequence 34, Appl

2	176	95.7	552	14	US-10-131-241-50	Sequence 50, Appl
3	131	71.2	632	14	US-10-131-241-51	Sequence 51, Appl
4	26	14.1	534	14	US-10-042-347-6	Sequence 6, Appl
5	26	14.1	537	14	US-10-131-241-59	Sequence 59, Appl
6	26	14.1	546	14	US-10-042-347-4	Sequence 4, Appl
7	26	14.1	549	12	US-10-252-418-3	Sequence 3, Appl
8	26	14.1	549	14	US-10-131-241-53	Sequence 53, Appl
9	26	14.1	551	13	US-10-080-797-2	Sequence 2, Appl
10	26	14.1	552	9	US-09-873-676-30	Sequence 30, Appl
11	26	14.1	574	14	US-10-060-036-64	Sequence 64, Appl
12	26	14.1	3394	10	US-09-880-107-2178	Sequence 1478, Ap
13	26	14.1	4551	14	US-10-080-036-144	Sequence 144, App
14	23	12.5	540	14	US-10-131-241-48	Sequence 48, Appl
15	23	12.5	552	12	US-10-292-418-17	Sequence 17, Appl
16	23	12.5	558	9	US-09-775-174-3	Sequence 3, Appl
17	23	12.5	558	9	US-09-775-325-3	Sequence 3, Appl
18	23	12.5	565	13	US-10-036-869-37	Sequence 37, Appl
19	23	12.5	573	10	US-09-988-831-12	Sequence 12, Appl
20	23	12.5	624	13	US-10-080-797-4	Sequence 4, Appl
21	20	10.9	900	14	US-10-131-241-45	Sequence 45, Appl
22	13	7.1	434	14	US-10-060-036-4	Sequence 4, Appl
C 23	9	4.9	1823	14	US-10-037-270-1012	Sequence 1012, Ap
C 24	9	4.9	3337	12	US-10-024-298A-169	Sequence 169, App
C 25	8	4.3	496	10	US-09-783-530-7378	Sequence 7378, Ap
C 26	8	4.3	624	13	US-10-027-632-229171	Sequence 229171,
C 27	8	4.3	624	13	US-10-027-632-229172	Sequence 229172,
C 28	8	4.3	696	13	US-10-027-632-21660	Sequence 21660, A
C 29	8	4.3	800	14	US-10-154-251-68	Sequence 68, Appl
C 30	8	4.3	800	14	US-10-154-251-69	Sequence 69, Appl
C 31	8	4.3	1080	9	US-09-925-301-294	Sequence 294, App
C 32	8	4.3	1518	14	US-10-156-761-2280	Sequence 2280, Ap
C 33	8	4.3	2274	14	US-10-156-761-5313	Sequence 5313, Ap
C 34	8	4.3	5798	11	US-09-993-525-1	Sequence 1, Appl
C 35	8	4.3	13046	9	US-09-764-870-595	Sequence 595, Appl
C 36	8	4.3	13046	11	US-09-764-891-5938	Sequence 5938, Ap
C 37	8	4.3	13046	14	US-10-125-540-595	Sequence 595, App
C 38	8	4.3	33719	15	US-10-080-170-651	Sequence 651, App
C 39	8	4.3	35937	10	US-09-782-378A-3	Sequence 3, Appl
C 40	8	4.3	111282	13	US-10-094-989-3	Sequence 3, Appl
C 41	8	4.3	9025608	14	US-10-156-761-1	Sequence 1, Appl
C 42	8	4.3	9025608	14	US-10-156-761-1	Sequence 1, Appl
C 43	7	3.8	29	13	US-10-080-797-15	Sequence 15, Appl
C 44	7	3.8	32	10	US-09-923-246-69	Sequence 69, Appl
C 45	7	3.8	32	14	US-10-295-723-69	Sequence 69, Appl

ALIGNMENTS

RESULT 1
US-10-292-418-34
; Sequence 34, Application US/10292418
; Publication No. US20030139365A1
; GENERAL INFORMATION:
; APPLICANT: Lo, Kin-Ming
; APPLICANT: Li, Yue
; APPLICANT: Gillies, Stephen D
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
; TITLE OF INVENTION: Immunofusins
; FILE REFERENCE: LEX-006C1
; CURRENT APPLICATION NUMBER: US/10/292,418
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/383,315
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: US 60/097,883
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS

; LOCATION: (1)..(552)
; OTHER INFORMATION: Endostatin
US-10-292-418-34

Alignment Scores: 1.01e-167 Length: 552
Pred. No.: 176.00 Matches: 176
Score: 176.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.65% Indels: 0
DB: 12 Gaps: 0

US-09-938-391-4 (1-184) x US-10-292-418-34 (1-552)

QY 9 ValLeuHisLeuValAlaLeuAsnSerProGlnProGlyGlyMetArgGlyLeuArgGly 28
DB 25 GTGCTGCACCTTGTGGCCCTTGAACACCGCCGACCGGGCGGATCGAGGCGATCCGGGGA 84
QY 29 AlaAspPheGlnCysPheGlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAla 48
DB 85 GCGGACTTCCAGTGTCTCCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 144
QY 49 PheLeuSerSerArgLeuGlnAspLeuTyrSerIleValArgAlaAspArgThrGly 68
DB 145 TTCTCTGCTGCGCGCTGCGAGGACCTCTACAGCATCGTGGCGCGCGCGCGCGCGCG 204
QY 69 ValProValValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSer 88
DB 205 GTGCCCGTGTCAACTCAGGAGCAGAGGTGCTCTTCCCGAGCTGGGAGGCTTATTCTCG 264
QY 89 GlySerGluGlyGlnLeuLysProGlyAlaArgIlePheSerPheAspGlyArgAspVal 108
DB 265 GGCTCCGAGGGCCAGCTGAAGCCCGGGCCCGCATCTTCTTTCGACGGCAGAGATGTC 324
QY 109 LeuGlnHisProAlaTrpProArgLysSerValTrpHisGlySerAspProSerGlyArg 128
DB 325 CTGCAGCAGCCCGCTGGCGCGCGGAGAGCGTGTGGCAGCGGCTCCGACCCAGCGGGCGC 384
QY 129 ArgLeuThrAspSerTyrCysGluThrTrpArgThrGluAlaProAlaThrGlyGln 148
DB 385 CGCCTGACCGACAGCTACTCGAGACGTGGCGACGAGGCGCCCGCGCGCGCGCGCG 444
QY 149 AlaSerSerLeuLeuAlaGlyArgLeuGlnGlnAlaAlaSerCysArgHisAla 168
DB 445 GCGTCTGCTGTGGCGGCGAGGCTGCTGGACAGAGGCGCGCGCGCGCGCGCGCGCC 504
QY 169 PheValValLeuCysIleGluAsnSerValMetThrSerPheSerLys 184
DB 505 TTCTGTGTGCTGTGCATCGAGAACAGCGTGCATGACCTCTCTTCTCCAAG 552

RESULT 2

US-10-131-241-50
; Sequence 50, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:

; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 552
; TYPE: DNA

; ORGANISM: Canine sp.
US-10-131-241-50

Alignment Scores: 1.01e-167 Length: 552
Pred. No.: 176.00 Matches: 176
Score: 176.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.65% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-4 (1-184) x US-10-131-241-50 (1-552)

QY 9 ValLeuHisLeuValAlaLeuAsnSerProGlnProGlyGlyMetArgGlyLeuArgGly 28
DB 25 GTGCTGCACCTTGTGGCCCTTGAACACCGCCGACCGGGCGGATCGAGGCGATCCGGGGA 84
QY 29 AlaAspPheGlnCysPheGlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAla 48
DB 85 GCGGACTTCCAGTGTCTCCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 144
QY 49 PheLeuSerSerArgLeuGlnAspLeuTyrSerIleValArgAlaAspArgThrGly 68
DB 145 TTCTCTGCTGCGCGCTGCGAGGACCTCTACAGCATCGTGGCGCGCGCGCGCGCGCG 204
QY 69 ValProValValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSer 88
DB 205 GTGCCCGTGTCAACTCAGGAGCAGAGGTGCTCTTCCCGAGCTGGGAGGCTTATTCTCG 264
QY 89 GlySerGluGlyGlnLeuLysProGlyAlaArgIlePheSerPheAspGlyArgAspVal 108
DB 265 GGCTCCGAGGGCCAGCTGAAGCCCGGGCCCGCATCTTCTTTCGACGGCAGAGATGTC 324
QY 109 LeuGlnHisProAlaTrpProArgLysSerValTrpHisGlySerAspProSerGlyArg 128
DB 325 CTGCAGCAGCCCGCTGGCGCGCGGAGAGCGTGTGGCAGCGGCTCCGACCCAGCGGGCGC 384
QY 129 ArgLeuThrAspSerTyrCysGluThrTrpArgThrGluAlaProAlaThrGlyGln 148
DB 385 CGCCTGACCGACAGCTACTCGAGACGTGGCGACGAGGCGCCCGCGCGCGCGCGCG 444
QY 149 AlaSerSerLeuLeuAlaGlyArgLeuGlnGlnAlaAlaSerCysArgHisAla 168
DB 445 GCGTCTGCTGTGGCGGCGAGGCTGCTGGACAGAGGCGCGCGCGCGCGCGCGCGCC 504
QY 169 PheValValLeuCysIleGluAsnSerValMetThrSerPheSerLys 184
DB 505 TTCTGTGTGCTGTGCATCGAGAACAGCGTGCATGACCTCTCTTCTCCAAG 552

RESULT 3

US-10-131-241-51
; Sequence 51, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:

; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
; LENGTH: 632
; TYPE: DNA
; ORGANISM: Murinae sp.

; CURRENT APPLICATION NUMBER: US/10/042,347
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: US 09/315,689
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 60/106,343
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 09/154,302
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 08/740,168
; PRIOR FILING DATE: 1996-10-22
; PRIOR APPLICATION NUMBER: US 60/005,835
; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-347-4

Alignment Scores:
Pred. No.: 5,52e-17 Length: 546
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-4 (1-184) x US-10-042-347-4 (1-546)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 121 GGGCTGGCGGGCACCTTCGCGCCCTCTCTGCTCGCGCTCGAGGACCTGTACAGCATC 180
QY 61 ValArgArgAlaAspArg 66
Db 181 GTGCGCGCTGCCGACCGC 198

RESULT 7

US-10-292-418-3
; Sequence 3, Application US/10292418
; Publication No. US20030139365A1
; GENERAL INFORMATION:
; APPLICANT: Lo, Kin-Ming
; APPLICANT: Li, Yue
; APPLICANT: Gillies, Stephen D
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
; FILE REFERENCE: LEX-006C1
; CURRENT APPLICATION NUMBER: US/10/292,418
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/383,315
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: US 60/097,883
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(549)
; OTHER INFORMATION: endostatin
US-10-292-418-3

Alignment Scores:
Pred. No.: 5,55e-17 Length: 549

Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 12 Gaps: 0

US-09-938-391-4 (1-184) x US-10-292-418-3 (1-549)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 121 GGGCTGGCGGGCACCTTCGCGCCCTCTCTGCTCGCGCTCGAGGACCTGTACAGCATC 180
QY 61 ValArgArgAlaAspArg 66
Db 181 GTGCGCGCTGCCGACCGC 198

RESULT 8

US-10-131-241-53
; Sequence 53, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-241-53

Alignment Scores:

Pred. No.: 5,55e-17 Length: 549
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-4 (1-184) x US-10-131-241-53 (1-549)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
Db 121 GGGCTGGCGGGCACCTTCGCGCCCTCTCTGCTCGCGCTCGAGGACCTGTACAGCATC 180
QY 61 ValArgArgAlaAspArg 66
Db 181 GTGCGCGCTGCCGACCGC 198

RESULT 9

US-10-080-797-2
; Sequence 2, Application US/10080797
; Publication No. US20020183253A1
; GENERAL INFORMATION:
; APPLICANT: Campochiaro, Peter A.
; APPLICANT: Dixon, Katharine H.
; APPLICANT: Brazzell, Romulus K.
; TITLE OF INVENTION: METHOD FOR TREATING OCULAR
; FILE REFERENCE: 4-31881A
; CURRENT APPLICATION NUMBER: US/10/080,797
; CURRENT FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Human
US-10-080-797-2

Alignment Scores:
Pred. No.: 5,57e-17 Length: 551
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 13 Gaps: 0

US-09-938-391-4 (1-184) x US-10-080-797-2 (1-551)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIle 60
Db 120 GGGCTGGCGGGCACCTTCGCGCTTCCTGCTCGCGCTGCAGGACCTGTACAGCATC 179
QY 61 ValArgArgAlaAspArg 66
Db 180 GTGCGCGCGTGCACCGC 197

RESULT 10

US-09-873-676-30

; Sequence 30, Application US/09873676
; Patent No. US20020077289A1
; GENERAL INFORMATION:
; APPLICANT: MacDonald, Nicholas J.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use
; FILE REFERENCE: 05213-0378 (43170-259333)
; CURRENT APPLICATION NUMBER: US/09/873,676
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-676-30

Alignment Scores:
Pred. No.: 5,57e-17 Length: 552
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 9 Gaps: 0

US-09-938-391-4 (1-184) x US-09-873-676-30 (1-552)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIle 60
Db 121 GGGCTGGCGGGCACCTTCGCGCTTCCTGCTCGCGCTGCAGGACCTGTACAGCATC 180
QY 61 ValArgArgAlaAspArg 66
Db 181 GTGCGCGCGTGCACCGC 198

RESULT 11

US-10-060-036-64
; Sequence 64, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.

; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-036-64

Alignment Scores:
Pred. No.: 5,77e-17 Length: 574
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-4 (1-184) x US-10-060-036-64 (1-574)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIle 60
Db 455 GGGCTGGCGGGCACCTTCGCGCTTCCTGCTCGCGCTGCAGGACCTGTACAGCATC 514
QY 61 ValArgArgAlaAspArg 66
Db 515 GTGCGCGCGTGCACCGC 532

RESULT 12

US-09-880-107-2178
; Sequence 2178, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2178
; LENGTH: 3394
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L22548
US-09-880-107-2178

Alignment Scores:
Pred. No.: 2,68e-16 Length: 3394
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
DB: 10 Gaps: 0

US-09-938-391-4 (1-184) x US-09-880-107-2178 (1-3394)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIle 60
Db 1624 GGGCTGGCGGGACCTTCGGCCCTTCCTCGCGCTGCAGGACCTGTACAGCATC 1683

QY 61 ValArgArgAlaAspArg 66
Db 1684 GTGCGCGGTGCGGACCGC 1701

RESULT 13

US-10-060-036-144
; Sequence 144, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:

; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566

; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144
; LENGTH: 4551
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-060-036-144

Alignment Scores:
Pred. No.: 3,456-16 Length: 4551
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.13% Indels: 0
Gaps: 0

US-09-938-391-4 (1-184) x US-10-060-036-144 (1-4551)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIle 60
Db 4120 GGGCTGGCGGGACCTTCGGCCCTTCCTCGCGCTGCAGGACCTGTACAGCATC 4179

QY 61 ValArgArgAlaAspArg 66
Db 4180 GTGCGCGGTGCGGACCGC 4197

RESULT 14

US-10-131-241-48
; Sequence 48, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:

; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22

; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 540
; TYPE: DNA

; ORGANISM: Rhesus monkey
US-10-131-241-48

Alignment Scores:

Pred. No.: 5,666-14 Length: 540
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.50% Indels: 0
Gaps: 0

US-09-938-391-4 (1-184) x US-10-131-241-48 (1-540)

QY 44 GlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIleValArgArg 63
Db 130 GGCACCTTCGGTCCCTTCCTCAGCGCTGCAGGACCTGTACAGCATCGTGCGCGGT 189

QY 64 AlaAspArg 66
Db 190 GCGACCGC 198

RESULT 15

US-10-292-418-17
; Sequence 17, Application US/10292418
; Publication No. US20030139365A1
; GENERAL INFORMATION:

; APPLICANT: Lo, Kin-Ming
; APPLICANT: Li, Yue
; APPLICANT: Gillies, Stephen D
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
; TITLE OF INVENTION: Immunofusins
; FILE REFERENCE: LEX-006C1
; CURRENT APPLICATION NUMBER: US/10/292,418
; CURRENT FILING DATE: 2002-11-12

; PRIOR APPLICATION NUMBER: 09/383,315
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: US 60/097,883
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(552)
; OTHER INFORMATION: endostatin

US-10-292-418-17

Alignment Scores:

Pred. No.: 5,776-14 Length: 552
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.50% Indels: 0
Gaps: 0

US-09-938-391-4 (1-184) x US-10-292-418-17 (1-552)

QY 44 GlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIleValArgArg 63
Db 130 GGCACCTTCGGGCTTCCTCCTAGGCTGCAGGATCTCTATAGCATCGTGCGCGGT 189

QY 64 AlaAspArg 66
Db 190 GCTGACCG 198

Search completed: August 18, 2003, 05:49:10
Job time : 738.833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 18, 2003, 00:28:06 ; Search time 1631.11 Seconds
(without alignments)
2741.704 Million cell updates/sec

Title: US-09-938-391-4
Perfect score: 184
Sequence: 1 HTHQDFQLVHLVALNSQP.....CRHAFVVLCIENSVMTSFSK 184

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Word size: 1

Total number of hits satisfying chosen parameters: 45562604

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -MODEL-frame-p2n.model -DEV-xlp
-Q/cgn2_1/USPTO.spool_p/US09938391/runat 04082003 130739 27598/app.query.fasta_1.718
-DB=EST -QFMT=fastap -SUFFIX=p2noli.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938391 @CEN 1 1 3399 @runat 04082003 130739 27598 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPOP=60 -XGAPOP=60
-FGAPEXT=7 -YGAPOP=60 -YGAPOP=60 -DELOP=60 -DELOP=60 -DELEXT=7

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pun.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_pbg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	44	23.9	451	10	BG383960	BG383960 302432 MA
2	44	23.9	491	10	BG383970	BG383970 302444 MA
3	43	23.4	279	9	AW464343	AW464343 BF230015B
4	43	23.4	551	10	BF074459	BF074459 221883 MA
5	43	23.4	623	14	CB424313	CB424313 598557 MA
6	43	23.4	657	14	CB444165	CB444165 695295 MA
7	43	23.4	709	14	CB443805	CB443805 594911 MA
8	37	20.1	387	10	BF601253	BF601253 266182 MA
9	32	17.4	483	12	BG823096	BG823096 602726433
10	27	14.7	577	12	BM488074	BM488074 pgm2n.pk0
11	27	14.7	614	13	BU459935	BU459935 603367327
12	27	14.7	652	13	BU352506	BU352506 603527982
13	27	14.7	678	14	CD215096	CD215096 pgm2n.pk0
14	27	14.7	686	13	BU290652	BU290652 604164521
15	27	14.7	706	13	BU439577	BU439577 603208252
16	27	14.7	717	13	BU230900	BU230900 603400730
17	27	14.7	736	13	BU268246	BU268246 603507092
18	27	14.7	755	13	BU262659	BU262659 603505164
19	27	14.7	756	13	BU380308	BU380308 603582387
20	27	14.7	776	13	BU485669	BU485669 603472003
21	26	14.1	390	9	AV660284	AV660284 AV660284
22	26	14.1	556	14	CB216999	CB216999 NISC nq11
23	26	14.1	616	13	BU158380	BU158380 AGENCOURT
24	26	14.1	634	13	BQ772348	BQ772348 UI-H-EZ1
25	26	14.1	639	12	BM998137	BM998137 UI-H-DT1
26	26	14.1	650	9	AI858615	AI858615 w140f01.x
27	26	14.1	660	9	AW192502	AW192502 x145610.x
28	26	14.1	664	13	BU632049	BU632049 UI-H-FE1
29	26	14.1	682	9	AW089583	AW089583 xg20f03.x
30	26	14.1	683	12	EM683067	EM683067 UI-E-E01
31	26	14.1	703	13	BU615520	BU615520 UI-H-FG0
32	26	14.1	707	10	BE908201	BE908201 601500458
33	26	14.1	715	9	AUI125614	AUI125614 AUI125614
34	26	14.1	757	10	BE906253	BE906253 601502237
35	26	14.1	832	10	BG387051	BG387051 602454749
36	26	14.1	846	13	BU540812	BU540812 AGENCOURT
37	26	14.1	881	14	CD105862	CD105862 AGENCOURT
38	26	14.1	884	12	BI161007	BI161007 602865213
39	26	14.1	929	13	BQ672290	BQ672290 AGENCOURT
40	26	14.1	944	13	BU859398	BU859398 AGENCOURT
41	26	14.1	947	13	BU556872	BU556872 AGENCOURT
42	26	14.1	979	13	BQ673186	BQ673186 AGENCOURT
43	26	14.1	1093	13	BQ723254	BQ723254 AGENCOURT
44	24	13.0	409	14	CB537919	CB537919 775852 MA
45	23	12.5	302	10	BF117397	BF117397 uz36e11.y

ALIGNMENTS

RESULT 1
BG383960
LOCUS BG383960 451 bp mRNA linear EST 12-MAR-2001
DEFINITION 302432 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BG383960
VERSION BG383960.1 GI:13308432
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 451)
AUTHORS Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,

Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R., Quackenbush
J. and Keele,J.W.
Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly

Mamm. Genome 13 (8), 475-478 (2002)

JOURNAL
MEDLINE
PubMed

COMMENT

EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
12226715
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACACGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCACGACG
Plate: 89 row: I column: 13
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers
1. 451
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC lPIG"
/note="Vector: pCMV SPORTS; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

FEATURES
source

1. 451
Location/Qualifiers

/organism="Sus scrofa"

/mol_type="mRNA"

/db_xref="taxon:9823"

/tissue_type="pooled"

/lab_host="DH10B"

/clone_lib="MARC lPIG"

/note="Vector: pCMV SPORTS; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

BASE COUNT 75 a 153 c 152 g 71 t
ORIGIN

Alignment Scores:
Pred. No.: 1.86e-31 Length: 451
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.91% Indels: 0
DB: 10 Gaps: 0

US-09-938-391-4 (1-184) x BG383960 (1-451)

QY 72 ValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGlu 91

Db 34 GTTAACCTCAGGACAGGAGTGCTGTCCCGAGCTGGAGGCGCTTGTCTCGGGCTCTGAG 93

QY 92 GlyGlnLeuLysProGlyAlaArgIlePheSerPheAspGlyValLeuGlnHis 111

Db 94 GGCCAGCTGAGCCGGCGCGCCGATCTCTTTTCAGCGGACAGACGTCCTTCAGCAC 153

QY 112 ProAlaTrpPro 115

Db 154 CCTGCTGGCCC 165

RESULT 2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Sus scrofa (pig)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

1 (bases 1 to 491)

Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R., Quackenbush
J. and Keele,J.W.
Porcine gene discovery by normalized cDNA-library sequencing and

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

JOURNAL

MEDLINE

PUBMED

COMMENT

EST cluster assembly

Mamm. Genome 13 (8), 475-478 (2002)

12226715

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 18

and -mismatch 12 options.

PCR Primers

FORWARD: AGGAACACGCTATGACCAT

BACKWARD: GTTTCCTCCAGTCACGACG

Plate: 89 row: K column: 13

Seq primer: ATTAGGTGACACTATAG.

Location/Qualifiers

1. 491

/organism="Sus scrofa"

/mol_type="mRNA"

/db_xref="taxon:9823"

/tissue_type="pooled"

/lab_host="DH10B"

/clone_lib="MARC lPIG"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

BASE COUNT 82 a 163 c 172 g 74 t

ORIGIN

Alignment Scores:

Pred. No.: 2.02e-31 Length: 491

Score: 44.00 Matches: 44

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 23.91% Indels: 0

DB: 10 Gaps: 0

US-09-938-391-4 (1-184) x BG383970 (1-491)

QY 72 ValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGlySerGlu 91

Db 34 GTTAACCTCAGGACAGGAGTGCTGTCCCGAGCTGGAGGCGCTTGTCTCGGGCTCTGAG 93

QY 92 GlyGlnLeuLysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeuGlnHis 111

Db 94 GGCCAGCTGAAAGCCGCGCGCCGATCTCTTTTCAGCGGACAGACGTCCTTCAGCAC 153

QY 112 ProAlaTrpPro 115

Db 154 CCTGCTGGCCC 165

RESULT 3

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Bos taurus

Bos taurus (cow)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos.

1 (bases 1 to 279)

Lewin,H.A., Soares,M.B., Rebeiz,M., Pardinas,J., Liu,L. and Larson
J.H.

Bovine ESTs

Unpublished

Contact: Lewin, H. A.

AW464343

BP230015B10C11 Soares normalized bovine placenta Bos taurus CDNA

clone BP230015B10C11 5', mRNA sequence.

AW464343

AW464343.1 GI:7034511

EST.

Bos taurus (cow)

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos.

1 (bases 1 to 279)

Lewin,H.A., Soares,M.B., Rebeiz,M., Pardinas,J., Liu,L. and Larson
J.H.

Bovine ESTs

Unpublished

Contact: Lewin, H. A.

W. M. Keck Center for Comparative and Functional Genomics
University of Illinois at Urbana-Champaign
340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL
61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu
Funding for cattle EST sequencing was provided by the USDA National
Research Initiative, Animal Genome Resource Grant AG 99-3205-8534
to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED
from Washington University Genome Center. Vector Trimmi g:
Cross_match from Washington University Genome Center PHRAP suite.
Sequences submitted are vector free and at least 200 bp in length.
PCR Primers
FORWARD: TAATACGACTCACTATAGG
BACKWARD: ATTAACCTCTCACTAAG
Insert Length: 279 Std Error: 0.00
Plate: BP230015B10 row: C column: 11
Seq primer: AGCGGATAACAATTTCACACGGA
High quality sequence stop: 279.

FEATURES

source
1..279
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="BP230015B10C11"
/sex="female"
/lab_host="DH108"
/clone_lib="Soares normalized bovine placenta"
/note="Organ: placenta; Vector: pVT3Pac; Site 1: EcoRI;
Site 2: NotI; The cDNA library was contributed by the
Soares laboratory and it was constructed and normalized
as described by Bonaldo, M.F., Lennon, G. and Soares,
M.B. (1996), Genome Research 6(9): 791-806."

BASE COUNT 35 a 106 c 86 g 52 t
ORIGIN

Alignment Scores:

Pred. No.: 9,99e-31 Length: 279
Score: 43.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.37% Indels: 0
DB: 9 Gaps: 0

US-09-938-391-4 (1-184) x AW464343 (1-279)

QY 70 ProValValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGly 89
DB 138 CCGCTCGTCAACCTCAGGACGAGGTGCTTTCTTAGCTGGAGGCTTGTCTCAGGC 197
QY 90 SerGluGlyGlnLeuLysProGlyAlaArgIlePheSerPheAspGlyArgAspValLeu 109
DB 198 TCGAGGGCCAGCTGAAGCCCGCCCGCATCTTCTCTTCAGCGCAGAGATGTCCTT 257
QY 110 GlnHisPro 112
DB 258 CAGCATCC 266

RESULT 4

BF074459
LOCUS BF074459 551 bp mRNA linear EST 25-APR-2001
DEFINITION 221883 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BF074459
VERSION BF074459.1 GI:10867970
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 551)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,

TITLE

JOURNAL
MEDLINE
PUBMED
COMMENT

Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-Mckown,C.G.,
Perteza,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keefe,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e.Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.

PCR Primers

FORWARD: AGGAACACAGCTATGACCAT

BACKWARD: GTTTCCAGTCACGACG

Plate: 81 row: F column: 9

Seq primer: ATTTAGCTGACACTATAG.

Location/Qualifiers
1..551

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/tissue_type="pooled"

/lab_host="DH108"

/clone_lib="MARC 2BOV"

/note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: Sall;
Library made from pooled tissue from testis, thymus,
semitendoneus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."

BASE COUNT 80 a 200 c 179 g 92 t
ORIGIN

Alignment Scores:

Pred. No.: 2e-30 Length: 551
Score: 43.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.37% Indels: 0
DB: 10 Gaps: 0

US-09-938-391-4 (1-184) x BF074459 (1-551)

QY 24 ArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaGlyLeuAla 43
DB 46 CCGCGCATCCGCGCGCGGACGCTTCAGTGCTTCCAGCAGCGCGCGGGCTGGCC 105
QY 44 GlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTySerIleValArgArg 63
DB 106 GGCACCTTCGCGCGGTTCCTGCTCCTCGCGGTTCGAGACCTGTACAGCATCGTGC 165
QY 64 AlaAspArg 66
DB 166 GCCGACCGT 174

RESULT 5

CB424313/c
LOCUS CB424313 623 bp mRNA linear EST 25-MAR-2003
DEFINITION 598557 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.
ACCESSION CB424313
VERSION CB424313.1 GI:29194232
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 623)

AUTHORS Smith, T.P.L., Roberts, A.J., Echtenkamp, S.E., Chitko-McKown, C.G.,
 Wray, J.E. and Keele, J.W.
TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL Unpublished
COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called with phred v0.020425.c and
 trimmed with the aid of the trim_alt option. Vector identified with
 cross_match v0.990329.
 Plate: FQY8006 row: I column: 12
 Seq primer: TAGAAGGCACAGTCGAGG.
 Location/Qualifiers
 1. 623

FEATURES

source
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 6BOV"
 /note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
 Library made with RNA pooled from multiple tissues
 including liver, lung, hypothalamus, pituitary, and
 placenta/endometrium."
 BASE COUNT 105 a 199 c 193 g 126 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2-27e-30 Length: 623
 Score: 43.00 Matches: 43
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 23.37% Indels: 0
 DB: 14 Gaps: 0
 US-09-938-391-4 (1-184) x CB424313 (1-623)

QY 70 ProValValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGly 89
 |||||
 Db 570 CCGTCTGCTCAACCTCAGGACGAGGTGCTTTCTTAGCTGGAGGCCCTGTCTCAGGC 511
 QY 90 SerGluGlyGlnLeuLysProGlyAlaArgGliePheSerPheAspGlyArgAspValLeu 109
 |||||
 Db 510 TCCGAGGGCCAGCTGAAGCCGCGCCGCGCATCTTCTCTTCGACGGCAGAGATGTCCTT 451
 QY 110 GlnHisPro 112
 |||||
 Db 450 CAGCATCCC 442

RESULT 6
 CB444165/c
 LOCUS 657 bp mRNA linear EST 25-MAR-2003
 DEFINITION 695295 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.
 ACCESSION CB444165
 VERSION CB444165.1 GI:29233914
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 657)
 Smith, T.P.L., Roberts, A.J., Echtenkamp, S.E., Chitko-McKown, C.G.,
 Wray, J.E. and Keele, J.W.
 A second set of bovine ESTs from pooled-tissue normalized libraries
 Unpublished
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366

Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called with phred v0.020425.c and
 trimmed with the aid of the trim_alt option. Vector identified with
 cross_match v0.990329.
 Plate: FQY8051 row: I column: 11
 Seq primer: TAGAAGGCACAGTCGAGG.
 Location/Qualifiers
 1. 657

FEATURES

source
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 6BOV"
 /note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
 Library made with RNA pooled from multiple tissues
 including liver, lung, hypothalamus, pituitary, and
 placenta/endometrium."
 BASE COUNT 107 a 217 c 219 g 113 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 2.4e-30 Length: 657
 Score: 43.00 Matches: 43
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 23.37% Indels: 0
 DB: 14 Gaps: 0

US-09-938-391-4 (1-184) x CB444165 (1-657)

QY 70 ProValValAsnLeuArgAspGluValLeuPheProSerTrpGluAlaLeuPheSerGly 89
 |||||
 Db 529 CCCGTCTGCTCAACCTCAGGACGAGGTGCTTTCTTAGCTGGAGGCCCTGTCTCAGGC 470
 QY 90 SerGluGlyGlnLeuLysProGlyAlaArgGliePheSerPheAspGlyArgAspValLeu 109
 |||||
 Db 469 TCCGAGGGCCAGCTGAAGCCGCGCCGCGCATCTTCTCTTCGACGGCAGAGATGTCCTT 410
 QY 110 GlnHisPro 112
 |||||
 Db 409 CAGCATCCC 401

RESULT 7
 CB443805
 LOCUS 709 bp mRNA linear EST 25-MAR-2003
 DEFINITION 694911 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION CB443805
 VERSION CB443805.1 GI:29233191
 KEYWORDS EST.
 SOURCE Bos taurus (cow)

ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 709)
 Smith, T.P.L., Roberts, A.J., Echtenkamp, S.E., Chitko-McKown, C.G.,
 Wray, J.E. and Keele, J.W.
 A second set of bovine ESTs from pooled-tissue normalized libraries
 Unpublished
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov

REFERENCE
 AUTHORS Smith, T.P.L., Roberts, A.J., Echtenkamp, S.E., Chitko-McKown, C.G.,
 Wray, J.E. and Keele, J.W.
 TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
 JOURNAL Unpublished
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called with phred v0.020425.c and
 trimmed with the aid of the trim_alt option. Vector identified with
 cross_match v0.990329.
 Plate: FQY8051 row: I column: 11
 Seq primer: GTAATACGACTCCTACTAGG.
 Location/Qualifiers

FEATURES

source

1..709
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"

/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."

BASE COUNT 96 a 288 c 218 g 107 t
ORIGIN

Alignment Scores:

Pred. No.: 2,59e-30 Length: 709
Score: 43.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.37% Indels: 0
DB: 14 Gaps: 0

US-09-938-391-4 (1-184) x CB443805 (1-709)

QY 24 ArgGlyIleArgGlyAlaAspPheGlnCysPheGlnGlnAlaArgAlaAlaGlyLeuAla 43
DB 559 CGCGCATCCGCGCGCGGACATTCAGTCTCCAGCAGCGCGCGGGGCTTGCC 618

QY 44 GlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTyrSerIleValArgArg 63
DB 619 GGCACCTTCGCGCGGTTCTGTCTCTCGCGGTTGCAGGACCTGTACAGCATCTGCGCGCG 678

QY 64 AlaAspArg 66

DB 679 GCCGACCGT 687

RESULT 8

BF601253

LOCUS BF601253 387 bp mRNA linear EST 25-APR-2001

DEFINITION 266182 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION BF601253

VERSION BF601253.1 GI:11698475

KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 387)

Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,
G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
Perteau,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keefe,J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA

libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

21180013

11282978

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 18

and -minmatch 12 options.

PCR primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCAGTCACGACG

Plate: 40 row: H column: 7

Seq primer: ATTTAGGTGACACTATAG.

FEATURES

source

Location/Qualifiers

1..387

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/tissue_type="pooled"

/lab_host="DH10B"

/clone_lib="MARC 3BOV"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendinosus muscle, and fetal
longissimus muscle."

BASE COUNT 52 a 140 c 131 g 64 t

ORIGIN

Alignment Scores:

Pred. No.: 6,46e-25 Length: 387
Score: 37.00 Matches: 37
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.11% Indels: 0
DB: 10 Gaps: 0

US-09-938-391-4 (1-184) x BF601253 (1-387)

QY 30 AspPheGlnCysPheGlnGlnAlaArgAlaAlaGlyLeuAlaGlyThrPheArgAlaPhe 49

DB 1 GACTTTCAGTCTTCAGCAGCGCGCGCGGGGCTGGCGGCACCTTCGCGCGGTC 60

QY 50 LeuSerSerArgLeuGlnAspLeuTyrSerIleValArgAlaAspArg 66

DB 61 CTGTCTCGCGTTTCAGGACCTGTACAGCATCGTGCAGCGCGCGCGCGT 111

RESULT 9

BF623096

LOCUS

DEFINITION

BF623096 483 bp mRNA linear EST 22-MAY-2001

602726433F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4865936 5',

mRNA sequence.

ACCESSION BF623096

VERSION BF623096.1 GI:14170683

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 483)

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: NIH Intramural Sequencing Center

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCW1732 row: a column: 09

High quality sequence stop: 479.

Location/Qualifiers

1..483

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4865936"

/tissue_type="adenocarcinoma cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_15"

/note="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"

BASE COUNT 69 a 200 c 146 g 68 t
ORIGIN

Alignment Scores:
Pred. No.: 4,26e-20 Length: 483
Score: 32.00 Matches: 45
Percent Similarity: 97.83% Conservative: 0
Best Local Similarity: 97.83% Mismatches: 0
Query Match: 17.39% Indels: 1
DB: 12 Gaps: 0

US-09-938-391-4 (1-184) x BG823096 (1-483)

QY 21 GlyGlyMetArgGlyLeuArgGlyAlaAspPheGlnCysPheGlnGlnAlaAla 40
DB 270 GCGCGCATCGGGGCATCGCGGGCGCGACTTCAGTGTTCAGCAGCGCGGGCGCT 329

QY 41 -GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTySeril 60
DB 330 GGGGCTGGCGGGCGACCTTCGCGCCTTCCTGCTCCGCGCTGCAGGACCTGTACAGCAT 389

QY 60 eValArgArgAlaAsp 65
DB 390 CGTGCCTGGTGGCGAC 405

RESULT 10

BM488074

LOCUS

DEFINITION

617 Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA clone pgm2n.pk006.h17 5' similar to gpAAC33294.2 (AF083440) collagen XVIII [Gallus gallus], mRNA sequence.

ACCESSION

BM488074

VERSION

BM488074.1 GI:18609005

SOURCE

Gallus gallus (chicken)

ORGANISM

REFERENCE

1 (bases 1 to 577)

AUTHORS

Cogburn, L.A. and Monsonego-Ornan, E.

TITLE

ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library, USDA/IRAFS Animal Genome Project

JOURNAL

Unpublished

CONTACT

Contact: Larry A. Cogburn

University of Delaware

Townsend Hall, Newark, DE 19717, USA

Tel: 302-831-1335

Fax: 302-831-2822

Email: cogburn@udel.edu, www.chickest.udel.edu.

Location/Qualifiers

1..577

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Commercial broiler and Ottawa Res. Centre

Strains 90 & 21"

/db_xref="taxon:9031"

/clone="pgm2n.pk006.h17"

/sex="Male and Female"

/tissue_type="Breast muscle, leg muscle and epiphyseal growth plate"

/dev_stage="Breast, leg: Embryo (d19) post-hatch (1d, 1.3, 5, 7, 9

11 weeks); growth plate (1d, 7d, 14d post-hatch)"

/lab_host="E. coli EM8H10B"

/clone_lib="Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm2n)"

/note="Vector: pCMVSPORT6; Library made from equivalent

pools of total RNA isolated from each tissue (embryonic muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth plate 33.3% of the final RNA pool). Single pass sequencing from 5'-end"

BASE COUNT 91 a 207 c 169 g 101 t 9 others
ORIGIN

Alignment Scores:

Pred. No.: 2,69e-15 Length: 577
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.67% Indels: 0
DB: 12 Gaps: 0

US-09-938-391-4 (1-184) x BM488074 (1-577)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerArgLeuGlnAspLeuTySeril 60

DB 259 GGGCTGGCGGGTACCTTCGTCCTCTCTCTCCGCTGCAGGACCTGTACAGCATC 318

QY 61 ValArgArgAlaAspArgThr 67

DB 319 GTGCGCAGGCGCGACCACT 339

RESULT 11

BU459935

LOCUS

DEFINITION

614 bp mRNA linear EST 29-NOV-2002
603367327F1 CSEQBNI9 Gallus gallus cDNA clone ChEST268f8 5', mRNA sequence.

ACCESSION

BU459935

VERSION

BU459935.1 GI:25949246

KEYWORDS

EST

SOURCE

Gallus gallus (chicken)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus

REFERENCE

1 (bases 1 to 614)

AUTHORS

Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,

Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.

TITLE

A Comprehensive Collection of Chicken cDNAs

JOURNAL

Curr. Biol. 12 (22), 1965-1969 (2002)

MEDLINE

22335534

PUBMED

12445392

COMMENT

Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology (UMIST

)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1..614

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Layer"

/db_xref="taxon:9031"

/clone="ChEST268f8"

/sex="Female"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="CSEQBNI9"

Vector: pBluescript II KS(+); Site_1:

EcoRI; Site_2: NotI; This normalized library was

constructed from 1 million independent clones. cDNA

synthesis was initiated using an oligo(dT) primer, using

methylated C in the first strand synthesis reaction.

Following this first strand reaction, double-stranded cDNA

was blunt-ended, ligated to NotI adapters, digested with EcoRI

, size-selected, and cloned into the NotI and EcoRI

compatible sites of a custom modified MCS of the

FEATURES

source

pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

BASE COUNT 106 a 212 c 179 g 117 t
ORIGIN

Alignment Scores: 2.87e-15 Length: 614
Pred. No.: 27.00 Matches: 27
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 14.67% Gaps: 0
DB: 13

US-09-938-391-4 (1-184) x BU459935 (1-614)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60
DB 183 GGGCTGGCGGTACCTTCGGTCTCTCTCCCGCTGCAGGACCTGTACAGCATC 242
QY 61 ValArgArgAlaAspArgThr 67
DB 243 GTGCGCAGGGCCGCGCACT 263

RESULT 12

LOCUS BU352506 652 bp mRNA linear EST 28-NOV-2002
DEFINITION 603527982F1 CSEQHN69 Gallus gallus cDNA clone CHEST477d22 5', mRNA sequence.

ACCESSION BU352506

VERSION BU352506.1 GI:25860507

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

1 (bases 1 to 652)

BOARDMAN, P.E., SANZ-EZQUERRO, J., OVERTON, I.M., BURT, D.W., BOSCH, E., FONG, W.T., TICKLE, C., BROWN, W.R.A., WILSON, S.A. and HUBBARD, S.J.

A Comprehensive Collection of Chicken CDNAs

Curr. Biol. 12 (22), 1965-1969 (2002)

22335534

12445392

CONTACT: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1. .652

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Compton Line 151"

/db_xref="taxon:9031"

/clone="CHEST477d22"

/sex="Female"

/tissue_type="cerebellum"

/dev_stage="adult"

/lab_host="PH10B"

/clone_lib="CSEQHN69"

/notes="Organ: brain; Vector: pBluescript II KS(+); Site 1:

EcORI; Site 2: NotI; This normalized library was

constructed from 1 million independent clones. cDNA

synthesis was initiated using an oligo(dT) primer, using

methylated C in the first strand synthesis reaction.

Following this first strand reaction, double-stranded cDNA

was blunt-ended, ligated to NotI adapters, digested with EcoRI

, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

BASE COUNT 116 a 232 c 185 g 119 t
ORIGIN

Alignment Scores: 3.05e-15 Length: 652
Pred. No.: 27.00 Matches: 27
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 14.67% Gaps: 0
DB: 13

US-09-938-391-4 (1-184) x BU352506 (1-652)

QY 41 GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle 60

DB 221 GGGCTGGCGGTACCTTCGGTCTCTCTCCCGCTGCAGGACCTGTACAGCATC 280

QY 61 ValArgArgAlaAspArgThr 67

DB 281 GTGCGCAGGGCCGCGCACT 301

RESULT 13

LOCUS CD215096

DEFINITION

CD215096

678 bp mRNA linear EST 20-MAY-2003

pgm2n.pk013.n20 Normalized chicken muscle cDNA library (pgm2n)

Gallus gallus cDNA clone pgm2n.pk013.n20 5' similar to

gb|AAC33294.2 (AF083440) collagen XVIII [Gallus gallus], mRNA

sequence.

CD215096

CD215096.1 GI:30954764

EST.

KEYWORDS

SOURCE

ORGANISM

Gallus gallus (chicken)

Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 678)

Cogburn, L.A. and Monsonego-Ornan, E.

Chicken ESTs from muscle

Unpublished

CONTACT: Larry A. Cogburn

University of Delaware

Townsend Hall, Newark, DE 19717, USA

Tel: 302-831-1335

Fax: 302-831-2822

Email: cogburn@udel.edu, www.chickest.udel.edu.

Location/Qualifiers

1. .678

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Commercial broiler chickens, Ottawa Research Ctr

Strains 90 & 21"

/db_xref="taxon:9031"

/clone="pgm2n.pk013.n20"

/sex="Male and Female"

/tissue_type="Breast muscle, leg muscle and epiphyseal

growth plate"

/dev_stage="Breast, leg: Embryo (dl9) post-hatch (ld.1,3,5,7,9

, 11 weeks); growth plate (ld.7d,14d post-hatch)"

/lab_host="E. Coli EMDH10B"

/clone_lib="Normalized chicken muscle cDNA library (pgm2n

)"

/note="Vector: pCMVSPORT6; Library made from equivalent

pools of total RNA isolated from each tissue (embryonic

muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth

plate 33.3% of the final RNA pool). Single pass sequencing

adapted from Soares et al., PNAS (1994) 91: 9228-9232 and
 Bonaldo et al., Genome Research 6 (1996): 791, except that
 a significantly longer reannealing hybridization was
 used."

BASE COUNT 125 a 253 c 202 g 126 t
 ORIGIN

Alignment Scores:

Pred. No.:	3,31e-15	Length:	706
Score:	27.00	Matches:	27
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	14.67%	Indels:	0
DB:	13	Gaps:	0

US-09-938-391-4 (1-184) x BU439577 (1-706)

QY	41	GlyLeuAlaGlyThrPheArgAlaPheLeuSerSerArgLeuGlnAspLeuTyrSerIle	60
Db	279	GGGCTGGCCGGTACCTTCGTCGCTTCCTCTCTCCCGCTGCAGGACCTGTACAGCATC	338
QY	61	ValArgArgAlaAspArgThr	67
Db	339	GTGGCGAGGGCCGACCACT	359

Search completed: August 18, 2003, 03:41:29
 Job time : 1633.11 secs

THIS PAGE BLANK (USPTO)